

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

REPRÉSENTATION ET SYMBOLE DANS LES SCIENCES COGNITIVES

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN PHILOSOPHIE

PAR
AHMED GHAZALI

JANVIER 2008

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

TABLE DES MATIÈRES

| | |
|--|----|
| LISTE DES FIGURES | iv |
| RÉSUMÉ | v |
| INTRODUCTION | 1 |
| PREMIÈRE PARTIE | |
| LA THÉORIE CLASSIQUE DE LA REPRÉSENTATION | 7 |
| CHAPITRE I | |
| LES FONDEMENTS THÉORIQUES DU MODÈLE CLASSIQUE | 8 |
| 1.1. Les origines de la représentation symbolique | 9 |
| 1.2. Du symbolisme au computationalisme | 12 |
| 1.3. Le fonctionnalisme computationnel, visite rapide | 14 |
| 1.4. Remarques générales | 18 |
| CHAPITRE II | |
| LA REPRÉSENTATION DANS LE MODÈLE CLASSIQUE | 21 |
| 2.1. Propriétés | 21 |
| 2.2. Critiques | 25 |
| DEUXIÈME PARTIE | |
| LA REPRÉSENTATION DANS LES MODÈLES ALTERNATIFS | 28 |
| CHAPITRE III | |
| LE CONNEXIONNISME ET LA REPRÉSENTATION DISTRIBUÉE | 29 |
| 3.1. Le réseau connexionniste | 29 |
| 3.2. Propriétés de la représentation dans le modèle connexionniste | 31 |
| 3.3. Diversité des réseaux connexionnistes | 36 |
| CHAPITRE IV | |
| LA THÉORIE DES SYSTÈMES DYNAMIQUES ET L'ANTI-REPRÉSENTATIONALISME | 38 |
| 4.1. La théorie des systèmes dynamiques | 38 |

| | |
|--|----|
| 4.2. Couplage et représentation | 40 |
| 4.3. Propriétés de la représentation dans les modèles dynamiques | 45 |
| TROISIÈME PARTIE | |
| APPROCHES INTEGRATIVES DE LA REPRÉSENTATION | 49 |
| CHAPITRE V | |
| VERS UNE REPRÉSENTATION MULTIFORME | 50 |
| 5.1. 'Remettre la représentation à sa place' | 50 |
| 5.2. Modèles hybrides | 55 |
| CHAPITRE VI | |
| LA DYNAMIQUE REPRÉSENTATIONNELLE CHEZ PIERCE | 61 |
| 6.1. L'œuvre de Pierce et la théorie dynamique | 61 |
| 6.2. La double fonction du concept de représentation | 63 |
| 6.3. La continuité du champ représentationnel | 67 |
| 6.4. L'évolution du processus représentationnel dans le temps | 71 |
| CONCLUSION | 74 |
| BIBLIOGRAPHIE | 80 |

LISTE DES FIGURES

| | | |
|-----------|--|----|
| Figure 1 | Succession chronologique des cadres théoriques qui ont fondé le cognitivisme | 18 |
| Figure 2. | Topographie des modèles cognitifs | 56 |

RÉSUMÉ

Le débat qui, au sein des sciences cognitives, a fait rage entre les défenseurs et les détracteurs de la représentation tend à l'apaisement. Les modèles connexionnistes et dynamiques n'ont pas réussi leur pari, celui de rendre caduc le concept de symbole et encore moins celui d'éliminer le concept de représentation du discours cognitif. Non seulement la représentation, comprise dans le sens général comme état médiateur porteur d'information, se maintient comme un concept utile et efficace pour décrire tous les systèmes cognitifs, mais la représentation symbolique reste encore le meilleur outil pour expliquer certains processus cognitifs de haut niveau. En revanche, les modèles connexionnistes et dynamiques ont réellement révolutionné le concept de représentation en le dotant de propriétés dynamiques (sensibilité au contexte, corporéité et temporalité inhérente) qui se sont imposées dans l'explication des processus cognitifs de niveau fin. Le débat se déplace. La question n'est plus de savoir quel est le paradigme gagnant mais celle de savoir quelle type de représentation convient pour décrire tel ou tel processus cognitif. La représentation se présenterait alors moins comme un concept figé que comme un cadre général ou un champ de potentiel qui se spécifie selon le processus cognitif à décrire. Cette approche est attestée par les modèles hybrides qui intègrent plusieurs formes représentationnelles au sein d'une même architecture. Elle trouve un support et une inspiration considérables dans l'œuvre de Pierce.

Mots-clés : représentation, symbole, fonctionnalisme, computationalisme, Connexionnisme, théorie des systèmes dynamiques, Pierce, signe.

INTRODUCTION

Depuis leur apparition, les sciences cognitives tentent de comprendre l'esprit, lieu de l'intelligence et de la pensée, en termes de son architecture. Orientée vers la structure et le mécanisme plutôt que vers la loi et la corrélation, leur démarche habituelle est de concevoir un modèle, le faire fonctionner puis observer sa performance traduite en terme de capacité à expliquer les phénomènes cognitifs réels. Sont ainsi nés différents courants qui tentent de décrire la relation de l'esprit avec le monde par des modèles variés et parfois contradictoires.

Un motif majeur de controverse et de divergence entre ces courants est le statut et le rôle de la représentation dans l'économie de la cognition. Le cognitivisme ou fonctionnalisme computationnel qui est l'hypothèse orthodoxe soutient que la représentation est de type symbolique. Son architecture pose le système cognitif comme un système de traitement de l'information qui manipule des symboles par le biais de règles syntaxiques. Cette conception qui a connu un succès et un essor considérables durant trois décennies fait depuis le milieu des années 80 l'objet de critiques et de remises en question.

Dans la mesure où la représentation est un concept central et fondamental au modèle classique, les critiques qui lui ont été adressées ont été particulièrement intenses. Il y a du vacarme dans la maison des sciences cognitives, et il serait trop simple de le réduire comme le font certains auteurs à un débat entre représentationalistes et anti-représentationalistes. La réalité est que les positions sont multiples et variées. Des modèles intègrent la représentation en lui retirant tout caractère symbolique, d'autres veulent bien parler de représentation symbolique à condition qu'elle ne soit pas régie par les règles de l'inférence logique. Enfin, il y a ceux qui défendent l'élimination

pure et simple du concept. Une même approche comme le connexionnisme nous donne à voir des réseaux localistes où la représentation se compare aisément à celle d'un modèle classique ainsi que des versions radicales non représentationnelles.

L'objet de cette recherche est de clarifier le débat autour du concept de représentation, en précisant le lien pas toujours clair qu'il entretient avec le concept de symbole auquel il a toujours été associé. La représentation se réduit-elle à un symbole comme le prétend les tenants du cognitivisme orthodoxe ? Dans le cas contraire et sachant que le symbole a une définition technique précise, comment peut-on définir le concept de représentation et comment l'articuler au symbole au sein de l'appareil cognitif ? Peut-on éliminer le concept et le discours de représentation comme le soutiennent certains courants de la théorie des systèmes dynamique ? Telles sont les questions que nous tenterons d'affronter. Notre allié dans cette entreprise est le temps. En effet, s'il y a encore du vacarme au sujet de la représentation, les signes de son épuisement sont là. Il semble que les belligérants ont dit ce qu'ils avaient à dire et qu'ils s'apprêtent maintenant à déposer les armes, et voire même à commencer les négociations ! Bref, il commence à y avoir du recul.

Notre démarche consiste dans un premier temps à déconstruire la théorie représentationnelle classique. Trois fondements théoriques contribuent à l'élaboration du concept de représentation dans le modèle classique : représentationnalisme, computationalisme et fonctionnalisme. La première version théorique du représentationnalisme remonte au moyen âge. La représentation y est comprise dans le sens général de re-présence de l'environnement dans l'esprit. Ce n'est qu'à partir de l'ère moderne sous la pulsion du cartésianisme que la représentation se trouve attribuée une nature symbolique. Le computationalisme est né des spéculations de Hobbes sur le raisonnement compris comme calcul symbolique. Suivent alors les étapes de la logique symbolique, de l'algèbre de Boole et des systèmes formels qui ont abouti à la première formulation de l'hypothèse computationaliste par Turing.

L'invention de l'ordinateur peu de temps après a fait le reste. Né d'une insatisfaction avec le néo-béhaviorisme et la théorie de l'identité, le fonctionnalisme élabore une théorie de l'esprit en exploitant les acquis du représentationalisme symbolique et du computationalisme.

Chacun de ces cadres théoriques dotera le concept de représentation d'un certain nombre de propriétés. Le représentationalisme lui attribue la propriété ontologique d'exister comme un état interne au système. Le représentationalisme symbolique lui confère la propriété sémantique d'être un symbole. Le computationalisme caractérise ce symbole comme un état discret compositionnel et une séquence dans un algorithme. Enfin le fonctionnalisme fait de la représentation un état intermédiaire, indépendant et abstrait.

Comprendre les choses de cette façon, permet de mieux apprécier et de mieux situer la critique qui sera adressée à la représentation. Si les modèles alternatifs pourront facilement récuser le fonctionnalisme, il ne pourront pas facilement se défaire du computationalisme et encore moins du représentationalisme qui s'avère une conception profondément ancrée dans le langage philosophique. Cela permet donc d'affiner le débat et de le sortir de la posture manichéenne de 'pour ou contre la représentation' dans lequel il s'enfonce parfois.

Deux théories alternatives se sont proposées de remplacer le cognitivisme comme modèle de la cognition : le connexionnisme et la théorie des systèmes dynamiques. Elles ont en commun un rejet du fonctionnalisme dont la structure représentationnelle est considérée comme un compromis fragile entre le cerveau et l'environnement ne rendant compte ni de l'un ni de l'autre. Le connexionnisme optera pour une représentation implémentée dans les structures neuronales du cerveau tandis que la théorie dynamique part de l'interaction entre l'agent et son environnement.

En concevant un réseau dont les éléments (les nœuds) sont porteurs de représentations, les réseaux connexionnistes sont largement représentationalistes. En considérant que le réseau calcule une fonction statistique sur la base des activations des nœuds, un grand nombre de modèles connexionnistes s'inscrivent dans l'hypothèse computationnel. Ce qui est rejeté, c'est le symbole et le calcul symbolique. Excepté dans les réseaux localistes, une représentation n'est pas encodée sur un seul nœud mais sur plusieurs. On dira qu'elle est distribuée. Cette architecture donne à la représentation connexionniste des propriétés dynamiques dont la tolérance à l'erreur et la sensibilité au contexte. En revanche, la représentation perd un peu de la stabilité qui est la marque d'une représentation symbolique. Pour cela, les réseaux connexionnistes ne pourront pas égaler les modèles classiques dans la description de certains phénomènes cognitifs comme la catégorisation ou la planification. Malgré tout, les modèles connexionnistes ont réussi à fournir une alternative sérieuse aux modèles classiques, et la représentation distribuée s'est imposée dans l'explication de nombreux processus cognitifs.

Tel n'est pas le cas de la théorie des systèmes dynamiques. Cette théorie élimine l'interface représentationnelle dans le lieu de l'interaction à l'intérieur de l'agent cognitif ou entre cet agent et son environnement. Si le couplage dynamique continue à être un centre d'intérêt, la thèse non représentationnelle perd de son éclat. Pour peu que l'on conçoive la représentation non pas dans le sens étroit de symbole mais dans le sens plus général d'un état médiateur porteur d'information, les couplages dynamiques sont aussi des représentations. La théorie poursuit la 'dynamisation' du modèle cognitif. Elle approfondit la sensibilité au contexte et la temporalité inhérente qui ont été initiées par le connexionnisme et rajoute un élément heuristique nouveau : la corporéité qui exploite le contact physique sensorimoteur entre l'agent cognitif et son environnement. Mais le caractère essentiellement transitoire de la représentation dans ces modèles rend la théorie inapte à décrire les phénomènes cognitifs qui ont des états stables. C'est le même défaut que dans le connexionnisme qui se trouve ici

aggravé. Par ailleurs, la difficulté de déterminer d'une manière précise et non arbitraire les paramètres individuels qui décrivent le système oblige la théorie dynamique à recourir à des paramètres collectifs, ce qui rajoute à l'instabilité physique de la représentation une crise sémantique. Tout cela a pour conséquence que la représentation dans ces modèles n'arrive pas au même niveau explicatif que la représentation symbolique et la représentation distribuée.

Les modèles connexionnistes et ceux de la théorie des systèmes dynamiques confèrent à la représentation des propriétés dynamiques incontestables qui se sont imposées dans l'explication de certains processus cognitifs. Néanmoins, ils ne détrônent pas, loin s'en faut, les propriétés symboliques (encodage, durabilité, stabilité et concaténation) qui demeurent les outils les plus efficaces pour expliquer certains processus cognitifs de haut niveau comme la catégorisation, le raisonnement analogique ou la planification. Mises côte à côte, toutes les propriétés se présentent par une série de couples bipolaires : discret/continu, durable/transitoire, abstrait/concret, etc. Cela laisse entrevoir la représentation comme un champ de potentialités dont les propriétés sont des continuums. Si notre intuition est fondée, le débat se déplace. Il ne s'agit plus de savoir si la représentation est seulement symbolique ou distribuée ou dynamique, ni de savoir lequel des modèles est le seul paradigme qui convient pour décrire la cognition. La question est plutôt de savoir comment la représentation peut prendre plusieurs formes et passer de l'un à l'autre parfois au sein du même processus cognitif.

Pourrait-on alors imaginer des mécanismes qui permettent d'expliquer ce continuum? Il nous semble que les recherches sur l'encodage et l'explicitation vont dans ce sens. Si le degré d'explicitation d'une représentation équivaut à la quantité d'effort à fournir pour l'extraction de l'information, et que cet effort dépend du contexte, nous avons là un mécanisme qui rend compte de la diversité des formes que peut prendre la représentation au sein d'un même processus cognitif. Dès lors, il ne devient plus

impossible d'envisager un modèle hybride qui intègre plus d'une forme représentationnelle au sein de son architecture. Nous présentons deux de ces modèles. Le premier, LISA de Holyoak et Hummel, est un modèle pour le raisonnement analogique qui intègre des représentations symboliques dans une architecture connexionniste. Le deuxième est le Système Symbolique Perceptuel de Prinz et Barsalou qui utilise un modèle classique où les représentations symboliques intègrent des propriétés dynamiques (sensibilité au contexte, corporéité et temporalité inhérente).

Cette approche processive et multiforme de la représentation n'est pas orpheline dans l'histoire de la philosophie. Elle a un parrain et non des moindres : Pierce. La catégoriologie peircienne définit la représentation à la fois comme *troisième* qui est la catégorie de la symbolisation et comme *signe triadique* qui est la totalité dynamique où se déploient simultanément les trois catégories. Dès son enclenchement, la représentation se manifeste d'emblée comme un processus de symbolisation. Mais ce processus peut ne pas stationner dans le symbole et 'dégénérer' dans des valeurs indexicales ou iconiques. Le champ représentationnel est en effet continu. Par ailleurs, le système de la représentation évolue dans le temps étant toujours orienté vers son attracteur qui est la généralisation en symbole. C'est alors la fin du processus et sa mort que seule l'intervention de la première catégorie, source de créativité et de chance, pourrait relancer de nouveau.

Une remarque est à faire avant de terminer cette introduction. Notre démarche est essentiellement transversale. Notre objectif est d'apprécier le statut et le rôle de la représentation dans les théories qui ont dominé les sciences cognitives, et non l'étude exhaustive de ces théories. C'est le concept de représentation qui est notre objet d'étude et non les modèles ou leurs auteurs.

PREMIÈRE PARTIE

LA THÉORIE CLASSIQUE DE LA REPRÉSENTATION

CHAPITRE I

LES FONDEMENTS THÉORIQUES DU MODÈLE CLASSIQUE

La théorie classique de la représentation fait partie intégrante de l'architecture conceptuelle du modèle classique de la cognition. Celui-ci reçoit plusieurs noms : Intelligence artificielle classique, symbolicisme, cognitivisme, fonctionnalisme computationnel, représentationnalisme computationnel, etc. Ce foisonnement de noms où se combinent à volonté les concepts et les *ismes* est pour le moins déconcertant. Il nous semble primordial de clarifier le rapport de la représentation avec chacun de ces concepts. Osons d'emblée une définition du modèle classique. Il est largement admis que ce modèle tient dans trois propositions :

1. Un processus cognitif est un système d'information. Conçu sur le modèle de l'ordinateur, il comporte deux niveaux indépendants de description, l'un physique d'implémentation et l'autre fonctionnel où se traitent les informations.
2. Ce traitement d'informations consiste en un processus de transformation qui est réglé selon un langage interne proche du langage formel de la logique.
3. Les informations traitées sont des représentations symboliques. Ce sont des états internes au système qui ont chacun un contenu qui renvoie à une entité externe.

Chacune de ces trois propositions renvoie à un cadre théorique précis. La première est une thèse du fonctionnalisme qui commande l'architecture du système cognitif. La deuxième est une thèse du computationalisme qui détermine la nature du traitement

informationnel. Enfin, la dernière est une formulation du représentationalisme symbolique qui identifie la nature de l'entité traitée. Tous les modèles du cognitivisme se réclament de ces trois propositions. Les trois propositions sont classées dans un ordre systématique. C'est d'abord l'architecture fonctionnaliste qui est donnée à voir, ensuite vient le processus computationnel et enfin la nature de l'entité traitée. Mais si on inverse l'ordre on trouve la succession chronologique de l'apparition de ces cadres théoriques. Représentationalisme symbolique, computationalisme et fonctionnalisme sont les trois moment-clés dans la longue histoire de la formation de la théorie classique de la représentation. En premier lieu, il y a le moment du symbolisme où la représentation acquiert l'identité de symbole notamment avec Descartes et Hobbes. Ensuite vient le moment de la computation où grâce à Turing le processus représentationnel est conçu comme un calcul algorithmique, thèse qui a lancé le programme de l'intelligence artificielle. Enfin, arrive le moment du fonctionnalisme computationnel où grâce à Putnam la représentation se trouve intégrée comme élément fondamental d'une théorie de l'esprit. On ne peut négliger cette perspective historique. Comme nous le verrons, elle permet de mieux situer les critiques qui seront adressées à la représentation ainsi que les modèles alternatifs qui seront proposés.

1.1. Les origines de la représentation symbolique

La représentation a une longue histoire comme ne manquent pas de le rappeler ses défenseurs :

There is no computation without representation. The representational account of the mind however predates considerably the invention of the computing machine. It is a throwback to classical epistemology, which is a tradition that includes philosophers

as divers as John Locke, David Hume, George Berkley, René Descartes, Immanuel Kant, John Stuart Mill and William James (Fodor, 1980)

Sur le plan théorique, l'intérêt philosophique pour le concept de représentation remonte au moyen âge. La tradition scholastique établit dans la ligne de l'aristotélisme un parallèle entre l'acte de connaître et l'acte de manger. Manger un objet, c'est absorber la matière de cet objet alors que connaître un objet c'est capturer sa forme pour en posséder une nouvelle présence en soi. La représentation est donc comprise comme re-présence. L'âme, à la différence du corps, ne peut pas traiter des entités physiques mais seulement des représentations. Parce qu'elle capture la forme, l'âme est informée par l'objet. Donc dès le moyen âge les concepts fondamentaux d'information et de représentation sont synonymes.

Quelle est la nature de la représentation ? Elle est surtout une ressemblance avec l'objet. Certes les médiévaux conçoivent bien que la représentation peut être aussi un symbole conventionnel comme c'est le cas des mots du langage naturel, mais les symboles linguistiques dépendent des pensées qui leur sont antérieures et qui sont des ressemblances. De cette façon les médiévaux résolvent le problème de la signification originelle. Le langage est symbolique mais il dépend de la pensée qui elle est iconique. La théorie de la ressemblance était privilégiée car elle était intuitive et se fondait sur la tradition antique du mimétisme. On peut en effet remonter l'idée à Platon. Le monde sensible est une copie, une imitation quoique imparfaite des Idées intelligibles. Le principe de la réminiscence est une re-présence des Idées. On doit aussi à Platon la distinction entre apparence et réalité et le fait que les deux sont liées par la ressemblance. L'apparent dépend du réel qui le conditionne à son image. C'est aussi l'idée chrétienne de l'homme fait à l'image de Dieu.

Avec la révolution copernicienne s'opère une séparation brutale entre apparence et réalité : le soleil nous apparaît tourner autour de la terre mais dans la réalité il en est

autrement. L'apparent n'est pas une imitation du réel mais une autre configuration radicalement différente. Il en résulte une conséquence importante : la réalité du monde est à concevoir indépendamment des apparences. On ne peut plus remonter de l'apparence de notre esprit à la réalité du monde, le principe de la représentation par ressemblance n'est pas satisfaisant et il convient de trouver d'autres types de représentations. C'est ce que Galilée va entreprendre. Si on ne peut pas se fier à nos apparences pour décrire le monde, on peut se fier aux mathématiques. Si l'apparent n'est pas une image fidèle du réel, il n'y a pas de raison que la géométrie ne soit qu'une image de l'espace. Ces deux constats conduisent Galilée à l'application des mathématiques aux phénomènes physiques et notamment aux problèmes du mouvement. Ainsi pour la première fois la nature physique est représentée par des symboles qui ne lui ressemblent pas : des rapports géométriques et arithmétiques entre des variables dénotant des quantités. C'est la naissance de la géométrie en tant que système de représentation symbolique indépendant qui tient par lui-même et qui s'applique à tel ou tel domaine de la réalité en dehors du principe de la ressemblance. Galilée introduit ainsi le principe de l'abstraction : les lignes ne représentent pas toujours des droites ou des distances mais des variables physiques : vitesses, durées, etc.

Cette révolution dans le domaine de la physique allait avoir des conséquences sur la représentation mentale. C'est à Descartes que revient le mérite de les faire ressortir en philosophie. Avec Descartes le mathématicien on assiste à un approfondissement de la représentation symbolique, et ce par le développement de la géométrie analytique. L'algébrisation de la géométrie a lieu et on peut désormais résoudre des problèmes géométriques par des méthodes algébriques. Descartes introduit la notation algébrique en équations et unifie ainsi l'algèbre et la géométrie. Une conséquence immédiate en fut l'algébrisation de la physique par une mise en équation des lois. Donc Descartes élève à un plus haut degré le procédé mental de l'abstraction : les

mathématiques ne traitent d'aucun sujet spécifique mais seulement des relations abstraites dans tel ou tel domaine.

Descartes le philosophe en tire une conséquence dans la théorie de l'esprit. Il applique ses découvertes sur les représentations en général au cas particulier de l'esprit. Les pensées sont des représentations symboliques analogues à celles employées en mathématiques. Il en résulte que le symbole est séparé de l'objet. De même que l'algèbre ne représente pas intrinsèquement des nombres, ce sont des notations différentes qui expriment des rapports numériques indépendamment de tout sujet en particulier, de même les représentations mentales sont indépendantes du monde et il n'y a pas de rapport intrinsèque entre la pensée et la chose. De même que la pensée des nombres ne dépend pas des nombres, la pensée du mot *Table* ne dépend pas de la chose *Table*. Les pensées sont d'un autre ordre symbolique et peuvent s'appliquer à tout autre objet dans le monde. Autrement dit, la pensée est un système de symboles dont il convient de trouver la logique indépendamment de ce qu'il représente. Ainsi est né le concept moderne de l'esprit, fondé sur le dualisme : la séparation radicale entre l'esprit et le monde, entre la pensée comme représentation symbolique et l'objet symbolisé.

1.2. Du symbolisme au computationalisme

En posant que l'esprit est une substance radicalement différente de la matière, le dualisme cartésien ne pouvait pas aller plus loin dans la précision de la nature du symbole mental et de la relation du symbole avec l'objet. Il fallait attendre le matérialisme de Hobbes pour commencer à se préoccuper de ces questions. «Par ratiocination, j'entends calcul» affirmait Hobbes dès 1650. Comme le langage naturel, la pensée est un discours mental qui consiste en des opérations symboliques. Elle diffère du langage uniquement par le fait qu'elle est un processus interne qui

s'exprime par des phantasmes (des parties de la pensée qui sont des marques cérébrales). La pensée suit des règles de calcul qui lui assurent clarté et rationalité. Hobbes a été inspiré par Galilée. Si pour ce dernier l'étude du mouvement n'est au fond que la géométrie, pour Hobbes la pensée n'est en soi qu'un calcul. De même que la géométrie est un système symbolique qui représente la réalité physique de la nature, notre pensée est un système de représentation symbolique. Si la réalité de l'univers est fondamentalement mathématique, il n'y a pas de raison que la réalité de la pensée ne le soit pas. Hobbes a aussi été inspiré par une invention de son époque : les machines à calculer de Schikard (1623) et de Pascal (1642). Ces machines effectuaient automatiquement les opérations arithmétiques de base. Hobbes établit un lien. Le raisonnement n'est jamais qu'un calcul dont la nature mécanique est confirmée par des constructions telles que la machine de Pascal. Le fait de matérialiser le calcul arithmétique n'est donc qu'une étape pour matérialiser le reste.

La naissance de la logique contemporaine intègre le concept hobbesien du raisonnement comme calcul. Dans *An investigation of the laws of thought* (1854) Boole accomplit l'algébrisation de la logique en substituant aux propositions les valeurs de vérité 0 ou 1. Devenant ainsi un calcul sur des nombres, la logique est mise sur la voie de la mécanisation comme auparavant l'arithmétique. Dans le même sens Frege (1879) construit une logique conçue comme un langage capable d'exprimer tout ce qui se fait et se dit en mathématiques. Il fonde ainsi la science des langages formels que développeront par la suite Russell et Hilbert.

Comme le note Haugeland (1986), si Hobbes est le grand père de l'IA, Turing en est le père. En effet, si avec Hobbes s'installe le principe de l'assimilation du raisonnement à un calcul, il restait à déterminer la nature de ce calcul. C'est véritablement Turing qui va accomplir cette mission. Turing part de l'analogie entre la machine et l'esprit humain. Il soutient qu'il y a des machines qui manipulent des symboles et que toute manipulation concrète peut être exécutée par une de ces

machines. D'où sa thèse : pour toute manipulation que l'esprit humain peut exécuter ou calculer concrètement et effectivement il existe une machine capable d'exécuter ce calcul. De là Turing propose une caractérisation exhaustive de tous les calculateurs symboliques possibles et montre qu'il existe une machine capable de simuler n'importe quelle autre machine. C'est la machine universelle qui fonde le concept de l'ordinateur.

En 1937 Shannon montre que les circuits électroniques peuvent être interprétés dans le calcul booléen établissant ainsi le lien entre électronique, calcul booléen, arithmétique et logique propositionnelle. Les formules propositionnelles et les opérations arithmétiques sont réalisables dans des circuits électroniques, ce qui conduit à la naissance de l'ordinateur.

La découverte de Turing initie toute la recherche de la philosophie de l'esprit contemporaine puisqu'elle permet de comprendre comment les phénomènes mentaux peuvent être matériellement réalisés. L'esprit comme l'ordinateur traite des informations représentées dans des structures de données ou symboles. Le contenu de la représentation symbolique n'est rien d'autre que le contenu de la pensée. Des règles s'appliquent à ces représentations symboliques et régissent leur transformation. Ces règles doivent être précises, explicites et sans exception, de sorte qu'un agent qui n'a pas idée de la signification des termes utilisés pourrait toujours exécuter le programme en suivant les instructions et en performant des opérations logiques et mathématiques élémentaires.

1.3. Le fonctionnalisme computationnel, visite rapide.

De Descartes à Turing les avancées de la représentation symbolique et du computationalisme s'appliquent au raisonnement et non à l'esprit ou les représentations mentales en général. C'est le domaine de la logique qui en était le

véritable champ d'application et l'antipsychologisme était de rigueur. Au mieux, la cybernétique fera en 1943 un lien avec la biologie en montrant que la logique propositionnelle est réalisée dans des neurones du cerveau et que l'esprit est matérialisé dans le tissu cérébral. Mais globalement, la vie mentale restait à la marge de ces recherches.

Or parallèlement à la prodigieuse histoire du calcul symbolique, se déroulait une autre histoire, celle de la philosophie de l'esprit. Menées de façon presque indépendante depuis plusieurs siècles, les deux histoires allaient se croiser dans les années soixante-dix. La première théorie sérieuse de l'esprit, on le sait, c'est le dualisme cartésien qui pose l'esprit comme une substance non matérielle. Très vite cette théorie devenait problématique et le monisme a été privilégié. Mise à part l'étoile filante du matérialisme de Hobbes, c'est surtout l'idéalisme qui a occupé la scène de Berkeley à l'idéalisme allemand en passant par Kant. Mais aussi impressionnant fût-il, l'idéalisme n'allait être qu'une parenthèse dans l'histoire de la philosophie de l'esprit telle que nous la concevons aujourd'hui et où le matérialisme s'est imposé comme le seul paradigme crédible. La question que se pose désormais les philosophes de l'esprit est : comment concilier le matérialisme devenu philosophie commune avec les phénomènes mentaux qui ont leur spécificité ? Autrement dit, comment concilier les faits de la psychologie empirique avec la philosophie de l'esprit ?

Le problème du dualisme est qu'il ne parvient pas à expliquer la causalité mentale. Comment une entité non matérielle peut-elle causer un comportement matériel sans violer les lois physiques de la conservation de la masse et de l'énergie ? Avec les progrès des sciences, la psychologie empirique appliquait de plus en plus les méthodes expérimentales utilisées dans les sciences de la nature, et il devenait difficile de trouver un argument qui parvienne à démontrer pourquoi la causalité esprit/corps ne peut être considérée comme un cas de causalité physique.

Le béhaviorisme logique a été la première alternative sérieuse au dualisme. Contrairement au behaviorisme radical qui considère que le comportement cognitif n'a pas de cause mentale et s'explique uniquement comme une réaction de l'agent à un stimulus, le béhaviorisme logique est une théorie sémantique qui considère que les états mentaux ont une signification. Attribuer un état mental à un organisme revient à dire que cet organisme est disposé à se comporter d'une certaine façon. La causalité mentale est ainsi reconnue mais elle est interprétée comme une manifestation de disposition au comportement. Le béhaviorisme logique rencontre plusieurs difficultés. La plus importante est qu'en traduisant les représentations mentales en dispositions au comportement, il postule implicitement qu'elles n'ont pas une existence propre. C'est pour pallier à cette difficulté que la théorie de l'identité est apparue comme une alternative. Cette théorie pose que les états et processus mentaux sont identiques aux états et processus neurophysiologiques dans le cerveau. Être dans un certain état mental équivaut à être dans un certain état neurophysiologique. La causalité mentale n'est rien d'autre que la causalité physique. La théorie de l'identité prend au sérieux le discours mental et considère que l'explication psychologique n'est pas seulement heuristique comme pour le béhaviorisme logique mais donne réellement l'histoire causale du comportement.

Ce que la théorie de l'identité introduit, c'est l'implémentation matérielle des processus mentaux jusque là ignorée. Mais en soutenant que les états mentaux qui peuvent exister ne peuvent être que neuronaux, la théorie prêche un physicalisme fort. Le mérite du béhaviorisme logique est d'avoir prouvé qu'il est possible de décrire les états mentaux par leurs relations causales et indépendamment de leur implémentation physique.

C'est ici qu'intervient le fonctionnalisme de Putnam (1960) dans un souci d'intégrer les points fort des deux théories précédentes. Le fonctionnalisme voudrait utiliser aussi bien le caractère causal des interactions entre le corps et l'esprit que le caractère

relationnel des propriétés mentales. Les états mentaux seraient alors individués par leur rôle causal dans la vie mentale de l'organisme et ne dépendraient pas de la matière qui les implémente. Quelle métaphore pourrait soutenir une telle architecture de l'esprit ? C'est ici que le fonctionnalisme exploite les acquis du représentationnalisme symbolique et du computationalisme. Le système cognitif fonctionne comme un ordinateur. C'est un système traitement de l'information. L'information est une représentation qui est d'une part concrètement implémentée dans les structures biologiques du cerveau et d'autre part elle a une nature sémantique en ce qu'elle renvoie à une entité dans l'environnement du système. Dès lors le processus d'un comportement cognitif peut être expliqué à deux niveaux d'abstractions indépendants qui correspondent à différentes disciplines : le niveau d'explication le plus concret est biologique, et le niveau le plus abstrait est fonctionnel. Expliquer une propriété mentale d'un comportement cognitif, c'est déterminer le processus de computation des états représentationnels responsables de ce comportement cognitif.

C'est ainsi que se parachève la théorie du fonctionnalisme computationnel qui est à la base de tous les modèles qui s'inscrivent dans le paradigme représentationnel classique. La cognition s'y présente comme l'opération d'un ordinateur mental situé dans le cerveau auquel les organes de sens livrent des représentations de l'état de son environnement. Le cerveau manipule cette information puis compute une spécification de l'action appropriée que le corps exécute. On présuppose donc que la représentation est une structure statique de symboles discrets et que l'opération cognitive est une transformation d'une structure statique à une autre. Ces transformations sont discrètes, instantanées et séquentielles. À la périphérie du système il y a les transducteurs qui transforment la stimulation sensorielle en des représentations input et des computations symboliques du cerveau en des représentations output.

1.4. Remarques générales

Chacune des trois théories s'est construite sur celles qui l'ont précédée. Il en résulte que le fonctionnalisme est un représentationalisme computationnel et que le computationalisme est un représentationalisme. Notons aussi que chaque théorie est largement indépendante de celles qui lui succèdent. Ainsi le représentationalisme est indépendant du computationalisme et du fonctionnalisme. De même, le computationalisme est indépendant du fonctionnalisme. Au contraire, le fonctionnalisme n'a aucune existence propre sans les concepts de représentation et de computation (Figure 1).

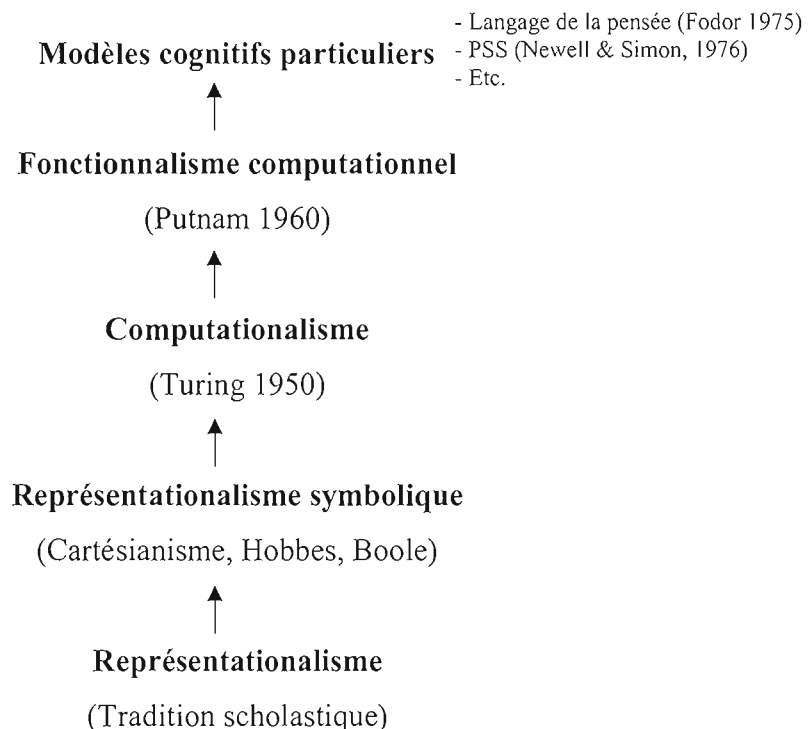


Figure 1. Succession chronologique des cadres théoriques qui ont fondé le cognitivisme
(d'après l'auteur)

Cette perspective historique est importante parce qu'elle permet de mieux cerner la portée des critiques et des modèles qui se proposeront pour remplacer le cognitivisme. Ainsi le connexionnisme ne rejettera que le cadre du fonctionnalisme et s'inscrira, du moins dans sa tendance majoritaire, dans un représentationnalisme computationnel. En revanche, la théorie des systèmes dynamiques projettera de rompre radicalement avec les trois fondements. Mais nous verrons que s'il sera facile aux auteurs de la théorie dynamique de démontrer une rupture avec le fonctionnalisme et le computationalisme, leur anti-représentationalisme sera une entreprise sans succès.

Chacun de ces cadres théoriques dotera le concept de représentation d'un certain nombre de propriétés. Le représentationnalisme lui attribue la propriété ontologique d'exister comme un état interne au système. Le représentationnalisme symbolique lui confère la propriété sémantique d'être un symbole. Le computationalisme caractérise ce symbole comme un état discret compositionnel et une séquence dans un algorithme. Enfin le fonctionnalisme computationnel fait de la représentation un état intermédiaire, indépendant et abstrait. Voir les choses de cette façon permet d'éviter de considérer la représentation comme un concept unitaire à prendre ou à laisser en bloc. Il faudrait plutôt voir dans la représentation un ensemble de propriétés qui découlent de différents cadres théoriques. Cela explique la diversité et les nuances des critiques qui lui seront adressées. Un modèle alternatif préservera certaines propriétés tandis qu'il rejettera d'autres pour les remplacer par de nouvelles.

La perspective historique de la représentation peut être interprétée d'une autre manière tout aussi intéressante. On peut dire que la théorie représentationnelle classique est née au confluent de deux histoires, celle du calcul symbolique, initiée par Descartes et Hobbes et culminant dans le computationalisme de Turing, et celle de la philosophie de l'esprit, initiée par le dualisme cartésien et qui se stabilise

efficacement dans le fonctionnalisme de Putnam. Cette interprétation nous donne à voir que la dynamique de cette histoire qui préside à la formation et à la transformation du concept de représentation ne s'arrêtera pas là. En effet, ce sont le développement des neurosciences et le progrès colossal des capacités de calcul informatique (notamment dans les circuits parallèles) d'une part, et l'insatisfaction croissante vis à vis du fonctionnalisme d'autre part, qui faciliteront la naissance de modèles alternatifs tels que le connexionnisme et la théorie des systèmes dynamiques.

CHAPITRE II

LA REPRÉSENTATION DANS LE MODÈLE CLASSIQUE

2.1. Propriétés

Le modèle classique confère à la représentation un nombre de propriétés que nous allons décliner. Nous nous efforcerons de les décrire et de les situer dans le cadre théorique qui les détermine. Plusieurs de ces propriétés ont été remarquées et rendues visibles à la lumière des critiques qui ont été adressées au concept. Ainsi la représentation est dite discrète par rapport à son caractère continu dans un modèle de la théorie dynamique et elle est locale par rapport à son caractère distribué dans un modèle connexionniste. Certaines propriétés semblent se recouper mais ces nuances auront une grande importance dans les modèles alternatifs. Si la qualité d'être discret implique celle d'être durable dans le modèle classique, dans un modèle connexionniste on rejettera le caractère discret tandis qu'on s'efforcera de produire des représentations assez durables afin de contester au fonctionnalisme le monopole de l'explication des états cognitifs qui requièrent des états stables.

- *La représentation est un état interne au système cognitif*

Cette propriété découle du représentationalisme dans sa version présymbolique. La représentation est la re-présence ou le substitut de l'environnement au sein du système. Nous le verrons, les modèles qui tenteront de se défaire de la représentation achopperont sur cette propriété.

- *La représentation est une entité durable*

La représentation dans le modèle classique se caractérise par sa stabilité dans le temps. Cette durabilité résulte du caractère symbolique de la représentation. Nous verrons que des modèles alternatifs tenteront de justifier une durabilité qui ne soit pas symbolique.

- *La représentation est une entité locale*

Cette propriété signifie qu'à chaque représentation atomique correspond un objet individuel dans le monde de sorte que le caractère sémantique de la représentation est clairement défini.

- *La représentation est un symbole*

Cette propriété découle du représentationalisme symbolique. C'est la propriété intentionnelle de la représentation, le fait qu'elle ait un contenu. On dit aussi qu'elle est au sujet de quelque chose d'autre.

Nous traitons avec des machines à traiter de l'information. Dans notre terminologie, ceci signifie que le mode de fonctionnement de ces machines consiste à utiliser des symboles qui tiennent la place des choses qu'ils représentent. (Marr 1982 : 21)

Le symbole se définit comme un élément qui est individué d'une manière non sémantique mais qui est le véhicule d'un certain type de contenu. L'analogie avec les langages naturels a été une démarche essentielle pour l'explicitation du symbole mental surtout dans la théorie du langage de la pensée.

To believe that such and such is to have a mental symbol that means that such and such tokened in your head in a certain way; it's to have such a token "in your belief box" as I'll sometimes say. (Fodor 1987, p.17)

- *La représentation est une entité discrète*

Tous les modèles classiques pose la représentation comme un élément discret. Les éléments d'une liste ou les mots d'une phrase sont discrets. Cette propriété permet que la même entité puisse être utilisée dans plusieurs combinaisons et à plusieurs moments. Par exemple, le processus de comparaison entre deux concepts implique la possibilité d'accéder à des entités discrètes reconnaissables et constantes dans le temps. Comme nous le verrons cette propriété sera durement critiquée parce qu'elle ne tient pas compte de la sensibilité de la représentation au contexte.

- *La représentation est un niveau indépendant*

Les représentations sont supportées par un système symbolique physique, qui est lui-même supporté par un niveau d'implémentation inférieure au sein d'un processeur d'un certain type physique qui, dans le cas de la cognition humaine, est le cerveau. Comme le niveau inférieur d'implémentation n'est pas nécessaire pour décrire exactement les phénomènes cognitifs, le niveau symbolique est donné comme exact et complet d'où son indépendance.

- *La représentation est abstraite*

C'est une propriété qui découle de l'architecture fonctionnaliste. La représentation entretient avec son contenu un lien sémantique mais elle en est physiquement indépendante ce qui la rend abstraite. Les modèles de l'action située s'insurgeront contre cette propriété et opteront pour des représentations concrètes qui naissent dans le contact physique entre le système cognitif et son environnement.

- *Les représentations ont une structure compositionnelle*

Le système symbolique est doté d'une syntaxe qui permet la compositionnalité sémantique de ses éléments. Les représentations atomiques mentales se combinent entre elles pour former des représentations moléculaires plus complexes. Ici aussi le

langage naturel sert comme modèle. En effet, les attitudes propositionnelles sont une juxtaposition d'unités qui représentent des objets (noms) avec des unités qui représentent des relations entre objets (verbes). Le contenu d'une représentation moléculaire est une fonction des contenus des représentations atomiques qui la constituent et des règles syntaxiques de combinaison.

In classical models, the principles by which mental states are transformed, or by which an input selects the corresponding output, are defined over structural properties of mental representations. Because classical mental representations have combinatorial structure, it is possible for classical mental operations to apply to them by reference to their form. (Fodor et Pylyshyn in Haugland 2000, p.317)

De cette propriété découle la systématique de la représentation ainsi que la productivité qui est la capacité de construire une infinité de représentations à partir d'un ensemble fini de symboles. Fodor et Pylyshyn (1988) ont défendu qu'aucun modèle cognitif ne peut rendre compte de la cognition s'il n'explique pas ces deux aspects. Les processus cognitifs comme la reconnaissance d'objets font appel à la compositionnalité de la représentation.

- *Les représentations sont régies par des règles de l'inférence logique*

Cette propriété découle de l'hypothèse computationaliste : un système est cognitif en vertu du fait qu'il calcule une fonction appropriée en exécutant un algorithme pour cette fonction. Il est donc essentiel que les représentations soient des objets de calcul manipulés par un algorithme.

What distinguishes what organisms do... is that a representation of the rules they follow constitutes one of the causal determinants of their behaviours (Fodor 1975, p.74)

Dans la théorie du langage de la pensée, on part de l'analogie avec le domaine linguistique. De même qu'on peut établir des algorithmes capables de produire des ensembles de phrases grammaticalement bien formées, dans la cognition on peut faire correspondre à une pensée une phrase d'un langage qui opère en réécrivant le pas à pas de la pensée par des symboles. Ce langage, *le mentalais*, désigne alors le système de codage des états mentaux qui sont structurés selon des règles formelles syntaxiques. Les règles qui opèrent dans ce langage relèvent de l'algèbre de Boole. Les processus cognitifs du raisonnement sont un exemple où cette propriété de l'inférence est efficacement utilisée.

2.2. Critiques

Pendant que le cognitivisme multipliait les modèles pour décrire différents processus cognitifs, grandissaient les doutes et les insatisfactions au sujet de son pouvoir explicatif. Ces remises en questions concernent en premier lieu le concept de représentation sur lequel se fondent tous les modèles classiques. Dans ce qui suit nous résumons les points majeurs de ces critiques.

a) Parce qu'elle est discrète, locale et durable, la représentation symbolique est fragile et rigide, ce qui la rend inapte à expliquer la plasticité et la robustesse de l'intelligence humaine. Comment l'esprit fonctionne face à des problèmes nouveaux ? Comment il réussit à s'en sortir avec un input dégradé ? D'où viennent les nouvelles représentations et les nouveaux concepts ? Comment le symbole se relie à son contenu ? Si on admet que des représentations discrètes et réglées conviennent à des processus cognitifs de haut niveau, qu'en est-il des structures fines de la cognition ?

b) L'architecture fonctionnaliste d'une représentation intermédiaire et indépendante ne peut pas rendre compte de l'interaction entre l'esprit et l'environnement. Le

modèle classique donne la priorité aux processus internes de la cognition au détriment des comportements sensoriels et moteurs. Sachant que les processus cognitifs sont souvent déclenchés par le contact avec l'environnement, n'est-ce pas par là que la conception du modèle doit commencer ?

c) La représentation fonctionnaliste ne nous dit rien sur son implémentation physique. Comment le cerveau donne naissance et porte le système physique symbolique ? En laissant de côté le niveau d'implémentation, n'a-t-on pas inversé les priorités ? Si les représentations mentales émergent des structures neuronales du cerveau alors logiquement la science cognitive devrait être la science de cette émergence.

d) Le caractère symbolique discret de la représentation ne tient pas compte de la continuité et de la trajectoire du développement humain. Le modèle classique est seulement capable de capter des états discrets à des moments donnés, mais cela ne donne pas l'histoire du développement.

De ces critiques sont nés deux grands courants qui ont proposé des modèles alternatifs. Le premier centre l'intérêt sur le cerveau tandis que le deuxième se tourne vers l'interaction avec l'environnement. Tout se passe comme si le 'compromis' fonctionnaliste éclate et l'on tente de renouer avec les modèles anciens. Là où le connexionnisme reprend le pari de la théorie de l'identité en implémentant la représentation dans les structures neuronales, la théorie des systèmes dynamiques semble renouer avec le projet behavioriste de traduire la cognition en matière de comportement face à un stimulus extérieur. Il ne s'agit pas là bien sûr d'une régression même s'il se trouve des auteurs pour le dire¹. En effet, entre les anciennes

¹ Nous renvoyons à ce sujet à l'article de Guy Tiberghien (in Rialle et Fiset, 1996) '*Le connexionnisme : stade suprême du behaviorisme*'. À noter que Tiberghien se réfère au connexionnisme radical qui défend la thèse anti-représentationaliste.

théories et les nouvelles il s'est écoulé un demi-siècle de progrès considérables dans le calcul, la simulation et les neurosciences.

DEUXIÈME PARTIE

LA REPRÉSENTATION DANS LES MODÈLES ALTERNATIFS DE LA COGNITION

CHAPITRE III

LE CONNEXIONNISME ET LA REPRÉSENTATION DISTRIBUÉE

3.1. Le réseau connexionniste

Le connexionnisme adopte une approche essentiellement empirique. Son intuition fondatrice est que la conception fonctionnaliste de l'esprit comme un ordinateur est erronée. Certes les états mentaux sont des représentations mais la structure porteuse de ces représentations mentales n'est pas du type d'ordinateur comme le cerveau est supposé être mais plutôt de type du réseau neuronal comme le cerveau se trouve réellement être. Il en résulte que le modèle qui convient le mieux pour décrire les phénomènes cognitifs est le réseau de neurones à l'image du cerveau qui est le modèle insurpassable.

Un réseau connexionniste est formé par des nœuds connectés. Chacun de ces nœuds se caractérise à tout instant par une variable qui représente son niveau d'activation et une constante qui représente son seuil. Quand l'activation du nœud dépasse son seuil, elle se propage depuis ce nœud vers les autres nœuds auquel il se trouve connecté. Chaque connexion a un poids qui détermine la proportion de l'activation qu'elle est en mesure de transmettre. Ce poids peut être positif ou négatif. Un nœud connecté négativement à un autre nœud l'inhibe en réduisant son activité. A tout instant, l'état du réseau est donné par la configuration des activations à cet instant et par les poids de toutes les connexions. L'entrée du système est l'état de la partie du réseau faite des nœuds qui sont désignés comme l'entrée. De même, la sortie du système est l'état de l'ensemble des nœuds désigné comme sortie quand le réseau se stabilise dans un état d'équilibre. On fournit un input au système en affectant aux nœuds d'entrées un

niveau d'activation. Cet input ainsi que la configuration du réseau (activations et poids) détermineront la propagation de l'activation à travers tout le réseau. Quand celui-ci se stabilise dans un état d'équilibre, l'état des nœuds qui constituent la sortie du système équivaut à l'output du réseau. Les lois du système sont données par des équations différentielles qui déterminent la propagation de l'activation en fonction du temps.

Un mot rapide sur la portée philosophique du connexionnisme qui a des conséquences profondes sur la conception de la représentation. Le fonctionnalisme computationnel se fonde sur le principe que l'observation d'un système cognitif permet de dégager des hypothèses décrivant la compétence de ce système et de produire ainsi cette compétence dans un modèle computationnel. Le modèle connexionniste soutient que les représentations internes qui entrent dans l'explication causale de la production du comportement cognitif ne sont pas celles qui interviennent dans la description commune. Autrement dit, on n'a pas accès à la compétence du système par le biais de l'intuition ou de l'introspection. Ainsi selon Dreyfus (1998), la description quasi-intuitive du comportement expert n'est pas une explication de ce comportement. Si pour le modèle classique, il est nécessaire d'avoir une théorie explicite sur la structure du problème à résoudre pour construire l'algorithme du modèle, pour le connexionnisme il n'est pas possible d'analyser les problèmes cognitifs complexes en raison de l'interaction non linéaire d'un grand nombre de facteurs. On reproche au représentationalisme computationnel de contourner le problème en simplifiant le système et en réduisant sa complexité au risque d'aboutir à un modèle abstrait qui, sur le plan pratique, est rigide et fragile.

Selon Dreyfus le modèle classique a tort de postuler que pour comprendre un domaine il faut posséder une théorie sur celui-ci. Le système expert du modèle computationnel ne peut pas rivaliser avec la maîtrise humaine. En effet, les systèmes experts suivent des règles alors que l'homme expert se sert des cas particuliers et

applique des heuristiques compliquées à des masses de faits. Dreyfus procède ainsi à une analyse phénoménologique de l'acquisition des savoir-faire. Qu'est-ce qu'un savoir-faire ? Et en quoi consiste l'expertise acquise par l'expert ? Il expose cinq stades de l'acquisition d'un savoir-faire et arrive à la conclusion que si le comportement du novice s'explique adéquatement avec le model de l'IA par le biais de représentations, règles et séquentialité, à partir du niveau de l'expertise il n'est plus possible d'utiliser les modèles symboliques computationels pour expliquer l'apprentissage. La raison est que l'expert n'applique pas des règles mais procède en discriminant des milliers de cas particuliers.

Cette approche détermine la forme et le processus représentationnels. Dans ce qui suit nous traiterons des propriétés de la représentation dans les réseaux connexionnistes. À noter que pour cette analyse nous ne considérons que les réseaux connexionnistes standards tels que ceux de Rumelhart (1989) et Smolensky (1989). Nous excluons les réseaux du connexionnisme radical tel que celui de Varela (1993) qui ne transforme pas un input en output par un processus représentationnel mais se présente comme un processus auto-organisateur, contraint par ses propres sorties et soumis à des modulations internes. Un tel réseau défend une thèse non représentationnelle qui le rend plus proche des modèles de la théorie des systèmes dynamiques qui seront traités dans le prochain chapitre.

3.2. Propriétés de la représentation dans le modèle connexionniste

- *La représentation est un état interne*

Cette propriété qui existait dans les modèles classiques se conserve. Le réseau connexionniste est une structure interne au système porteuse des représentations mentales. Le connexionnisme est donc un représentationalisme.

- *La représentation est implémentée*

Les représentations mentales sont considérées comme des émergences du niveau d'implémentation inférieur.

Emergent properties are high-level effects that depend on lower-level phenomena in some systematic way (Churchland and Sejnowski in Eliasmith 1996, p.304)

Cette propriété a pour vocation de s'opposer à la propriété de l'indépendance de la représentation symbolique par rapport à ce qui l'implémente dans le cerveau. Pour rendre compte de l'implémentation, le réseau connexionniste n'élimine pas pour autant le niveau représentationnel mais rajoute un niveau inférieur (subsymbolique) où ces représentations se trouvent formalisées dans une structure neuronale sur le modèle du cerveau. Smolensky insiste sur cette double signification du réseau :

Les deux niveaux sont essentiels : le niveau inférieur pour définir ce qu'*est* le système (en terme de propagation d'activation), et le niveau supérieur pour comprendre ce que le système *signifie* (dans les termes du domaine du problème.) (Smolensky, in Andler 1993, p.86)

Mais s'il garde le niveau des représentations macroscopiques, il lui retire la capacité d'explication exacte, les structures mentales n'interviennent que dans les descriptions approximatives. Cette formalisation n'est pas sans poser des problèmes. Si le niveau subsymbolique est intermédiaire entre le cerveau et le niveau des représentations macroscopiques peut-on parler d'une réelle implémentation ? Les auteurs du connexionnisme eux-mêmes insistent sur le fait de ne pas confondre le fonctionnement de leurs réseaux avec le fonctionnement neuronal du cerveau qui demeure encore insuffisamment connu. Dans ce cas ne serait-il pas plus heuristique de considérer que la cognition fonctionne avec des processus susceptibles d'une description indépendante de la biologie du cerveau ? Fodor et Pylyshyn (1988) citent

l'exemple de la théorie de la tectonique des plaques qui est une description exacte et complète des mouvements continentaux, indépendante de la physique microscopique.

- *La représentation est distribuée*

La représentation est portée par des nœuds du réseau. Dans les réseaux localistes, chaque nœud représente un concept spécifique et fonctionne donc comme un symbole. Un exemple en est le réseau où les nœuds représentent les membres d'un groupe. Si deux membres sont en relation alors la connexion entre eux est positive, sinon elle est négative. Ce réseau simple est un reconnaissseur automatique de relation. Si on active un nœud alors tous les nœuds qui sont en relation avec lui sont activés. C'est un réseau sémantique, chaque nœud a une interprétation spécifique et représente quelque chose de fixe. Dans ce cas la représentation est symbolique comme dans les modèles classiques.

Mais les réseaux localistes ne sont qu'un cas particulier exceptionnel. Dans la majorité des cas, les réseaux connexionnistes sont plus complexes et il n'existe pas de relation spécifique entre un nœud et un concept particulier. La représentation est portée par plusieurs nœuds. Elle est dite distribuée. Un exemple classique en est le perceptron multi-stratifié qu'on utilise pour des tâches de classification. Un tel réseau a une grande quantité de nœuds dans la strate cachée, chacun de ces nœuds étant connecté avec les nœuds de la strate input. L'information qui arrive dans le réseau est distribuée sur tous les nœuds cachés, de sorte qu'aucun de ces nœuds ne représente quelque chose de spécifique. Les nœuds n'ont ainsi aucune interprétation particulière. Ils n'ont pas de caractérisation sémantique qui les définisse individuellement. Aucun nœud ou poids n'a aucune signification propre. Les nœuds et les poids participent toutes les fois que le réseau est en marche indépendamment de toute signification particulière. Le degré de distribution dépend de la complexité du problème.

Les réseaux distribués ont une capacité à la généralisation supérieure aux modèles de la représentation symbolique locale. Les réseaux distribués sont aussi solides et ont une grande tolérance à l'erreur dans la mesure où tous les nœuds participent et que l'efficacité du réseau ne dépend pas d'un nœud en particulier. Dans le cas de la représentation symbolique, un symbole qui ne s'active pas en raison d'une panne peut provoquer des dégâts dans la computation en raison de la propriété sériale. En revanche, le parallélisme d'un réseau connexionniste fait qu'un nœud endommagé n'a pas beaucoup d'effet sur le réseau. La raison est qu'aucune propriété du réseau n'est reliée à une propriété de nœud. Un autre avantage important de la représentation distribuée est la sensibilité au contexte. Comme la représentation correspond à des vecteurs dans l'espace inférieur subsymbolique, certains de ces vecteurs renvoient aux particularismes du contexte tandis que d'autres représentent l'invariant de l'état cognitif.

L'argument le plus célèbre contre la représentation distribuée est de celui de Fodor et Pylyshyn (1988). Ces auteurs commencent par insister sur la systématisme comme caractéristique fondamentale des processus cognitifs.

Cognitive capacities come in structurally related clusters; their systematicity is pervasive. (Fodor et Pylyshyn in Haugeland 2000, p. 334)

Mais la représentation distribuée n'a pas une structure syntaxique qui lui permettrait de rendre compte de cette systématisme et l'expliquer.

Because it [connectionist architecture] acknowledges neither syntactic nor semantic structure in mental representations, it perforce treats them not as generated set but as a list. But lists, qua lists, have no structure. (p. 334)

Cela rend la capacité explicative de la représentation distribuée inférieure à celle de la représentation symbolique dans une large gamme de la cognition qui fait appel à la systématique et qui contient des processus aussi divers que la catégorisation, le raisonnement, l'utilisation du langage, la reconnaissance de scènes ou d'objets complexe, etc.

- *La représentation est régie par des lois dynamiques*

Les représentations distribuées ne se combinent pas par concaténation comme c'est le cas pour les symboles dans le modèle classique. Les lois qui régissent le système sont les lois de propagation des activations et les lois de modification des poids de connexions. Ces lois se traduisent en équations différentielles. Les règles de propagation sont des règles d'inférence statistique (et non logique). Les règles de modification des poids sont des procédures de mise en mémoire et d'apprentissage. Dans ce sens, le connexionnisme est une variante du computationalisme. Ce qui le sépare du modèle classique c'est que ce dernier suppose que les symboles représentés sont manipulés par des règles de la logique formelle. Ce n'est donc pas le principe de la computation qui est en cause, mais plutôt ses modalités : inférence logique dans le cas du cognitivisme et généralisation statistique dans le cas du connexionnisme.

Parce que l'activation est une variable dynamique continue, la représentation est transitoire. Quand une configuration d'activations s'installe, l'ancienne configuration qui lui précédait disparaît aussitôt pour laisser place à la nouvelle. Si ce caractère transitoire permet au réseau connexionniste de décrire des processus cognitifs de bas niveau, il le pénalise pour les processus cognitifs dont les représentations sont des états stables comme le raisonnement ou la formation des catégories. C'est une réelle difficulté que les réseaux connexionnistes tentent de minimiser en justifiant un minimum de stabilité et de durabilité dans leurs représentations distribuées. Ils

essaient pour cela de recourir à des discontinuités dynamiques comme le seuil d'activation, l'harmonie ou l'attracteur.

La catégorisation ne consiste pas dans l'exécution d'un algorithme symbolique, mais dans l'évolution continue du système dynamique, une évolution qui conduit les états vers des attracteurs, vers une harmonie maximale. (Smolensky in Andler 1993, p. 100)

3.3. Diversité des réseaux connexionnistes

La représentation dans les réseaux connexionnistes se caractérise par une grande diversité selon le type de réseau. Comme le soutient Tiberghien, il serait plus judicieux de parler des connexionnistes plutôt que de connexionnisme.

Les différences entre ces modèles connexionnistes, ou néo-connexionnistes, sont au moins aussi importantes, sinon plus, que leur similitudes. Traces composites ou séparées, localisées ou distribuées; traces stockées sous forme de vecteurs, de matrices 2D ou 3D; récupération par échantillonnage, résonance, association, corrélation. De façon plus générale ces différences portent donc à la fois sur le codage de l'information, son mode exact de stockage, les mécanismes de récupération et, enfin, l'architecture générale du système de mémoire. (Tiberghien in Riaille et Fiset 1996, p.27)

Certains auteurs comme van Gelder (1997) ont soutenu que le connexionnisme serait une conception hybride entre le computationalisme et le dynamisme parce qu'il emprunte ses modèles tantôt à l'un tantôt à l'autre. Cette situation instable du connexionnisme entre deux paradigmes qui ont une plus grande cohérence théorique serait, selon van Gelder, un indice qu'il est une étape dans l'évolution qui va de l'approche computationaliste à l'approche dynamique de la cognition. Le

connexionnisme préparerait donc le terrain pour la théorie dynamique avant de s'éclipser pour de bon. Cette hypothèse semble être de plus en plus infirmée par les récents développements. Au contraire, il semble que la 'plasticité' du connexionnisme et sa capacité à intégrer tant des représentations symboliques locales que des mécanismes dynamiques, le destineraient à jouer un rôle très important. Comme nous le verrons au sixième chapitre, le débat sur la représentation semble vouloir se résoudre dans l'exploration de modèles hybrides qui intègrent plusieurs formes de représentations. Dans ce cas le connexionnisme se trouverait à être le modèle le plus favorisé. Si bien qu'il existent des auteurs qui retournent l'argument de van Gelder contre lui : ce n'est pas le connexionnisme qui préparerait le terrain pour la théorie dynamique, mais au contraire, c'est la théorie dynamique qui travaillerait pour le compte du connexionnisme puisque ce dernier ne cesse d'intégrer dans ses réseaux la recherche dynamique de la première. Car s'il est vrai que la théorie dynamique a plus de cohérence conceptuelle, elle est plus limitée dans ses capacités explicatives. De cela nous traiterons dans le chapitre qui suit.

CHAPITRE IV

LA THÉORIE DES SYSTÈMES DYNAMIQUES ET L'ANTI- REPRÉSENTATIONALISME

4.1. La théorie des systèmes dynamiques

Les auteurs qui se réclament de la théorie des systèmes dynamiques ou dynamicisme comme le nomment certains (Eliasmith 1996), soutiennent que la cognition n'est ni une manipulation de symboles ni un traitement connexionniste mais une interaction complexe et dynamique entre un agent cognitif et son environnement. Pour cela, ils annoncent une rupture radicale avec le cognitivisme et le connexionnisme. Selon eux, la cognition a une dynamique propre dont le meilleur modèle est celui de la dynamique des systèmes largement utilisé dans les sciences de la nature. L'esprit et l'environnement sont couplés directement au sein d'un même système, ce qui rend inutile et inefficace toute interface représentationnelle entre les deux.

L'idée de l'esprit comme système dynamique peut être remontée à David Hume mais l'approche contemporaine est datée du début de la cybernétique (Ashby, 1952). Dans les décades suivantes la recherche sur la dynamique a été menée dans des programmes aussi divers que la psychologie écologique, la morphodynamique, la synergetique et la neurologie. Dans les années 80, l'insatisfaction croissante à l'égard du fonctionnalisme computationnel, le développement des mathématiques des systèmes non linéaires ainsi que les avancées informatiques dans le domaine de la simulation ont été les facteurs qui ont contribué à l'éclosion des recherches sur la dynamique cognitive. A partir des années 90, plusieurs auteurs soutiennent que

l'approche dynamique a suffisamment de scope et de cohérence pour compter comme un paradigme original et indépendant.

Natural cognitive systems are certain kinds of dynamical systems, and are best understood from the perspective of dynamics. (van Gelder in Port et van Gelder, 1995, p.4)

Un modèle dynamique décrit le comportement cognitif de l'agent par des concepts géométriques : espace d'état, attracteur, trajectoires, et topologie. L'espace d'état du système est l'espace défini par l'ensemble de tous les états possibles par lesquels passe ou pourrait passer le système. La trajectoire est une succession particulière de plusieurs états du système et elle est l'expression de son comportement. La topologie de l'espace d'état est la description de tous les points de cet espace. Un attracteur est un point ou un chemin dans l'espace d'état vers lequel la trajectoire va tendre quand elle est dans le voisinage de cet attracteur. Le modèle décrit le comportement du système en utilisant des paramètres ou variables interprétés. Le nombre de ces paramètres est le nombre des dimensions du modèle. La théorie préconise un nombre faible de paramètres. Et elle s'oppose en cela au réseau connexionniste qui affiche une dimensionnalité très élevée, chaque nœud comptant comme une dimension. Le système est gouverné par des équations différentielles non linéaires couplées qui consistent en des constantes et des paramètres. Par exemple, le mouvement d'une pendule verticale est décrit par une équation à une variable qui est l'angle avec la verticale et une constante qui est la constante du champ gravitationnel.

La théorie des systèmes dynamiques fournit le cadre théorique mathématique qui décrit comment les processus changent dans le temps. C'est le modèle le plus utilisé et le plus efficace dans les sciences naturelles. L'équation différentielle décrit comment le système cognitif évolue dans le temps, c'est-à-dire comment le changement d'une de ses variables d'état dépend à tout instant de la valeur de cette

variable ainsi que des autres paramètres du système. Par conséquent, alors que pour le computationalisme, les règles qui gouvernent le comportement du système sont définies en fonction des entités qui ont un statut représentationnel (symboles ou nœuds), pour la théorie des systèmes dynamiques les règles sont définies en fonction des états numériques.

La théorie des systèmes dynamiques rejette le postulat cartésien d'un cerveau inné et encapsulé et considère le système cognitif comme une totalité cerveau-corps-environnement. Autrement dit, le système cognitif n'entre pas en interaction avec le monde en passant des messages et commandes, il coévolue avec lui. Par conséquent, le comportement de l'agent humain ne peut pas être expliqué par les ressources de la représentation. Les écrits des dynamicistes regorgent de sentences condamnatoires à l'égard de la représentation.

It is the concept of representation which is insufficiently sophisticated. (van Gelder, 1993 : 6).

Representation is the wrong unit of abstraction in building the bulkiest parts of intelligent systems. (Brooks, 1997: 396).

We are not building representations at all! (Thelen et Smith, 1994 : 338).

4.2. Couplage et représentation

La théorie dynamique remplace la représentation par le couplage entre les paramètres du système. Le couplage implique que le changement dans un paramètre provoque le changement des autres paramètres couplés. On cite souvent comme exemple les robots de Brooks (1991) qui ne forment pas des structures représentationnelles

compliquées pour décrire leurs environnements mais utilisent uniquement l'information qui est immédiatement disponible et ne stockent que des informations transitoires pendant qu'ils naviguent dans le monde. Les modules du robot communiquent entre eux en agissant l'un sur l'autre par exemple par une action d'inhibition du mouvement, mais pas en passant des représentations de l'un à l'autre.

On doit à van Gelder (1997) l'un des arguments les plus célèbres en faveur du couplage non représentationnel. L'auteur utilise pour son argument l'exemple du régulateur dynamique de Watt. Le problème se pose comme suit : la valve du piston à vapeur doit avoir une ouverture réglable au temps précis pour compenser les variations de la pression de vapeur et maintenir de cette façon relativement constante la vitesse de la rotation de la roue. Comment réaliser techniquement cette solution? Un axe vertical ancré dans la roue épouse le même mouvement de rotation. Attachés à cet axe, deux bras ont à la fin de chacun une boule métallique. Quand l'axe tourne, la force centrifuge entraîne les boules vers l'extérieur puis vers le haut. Ce mouvement centrifuge des bras est lié à la valve du piston. Il en résulte que quand la vitesse de la roue est excessive les bras montent et cela entraîne une diminution de l'ouverture de la valve et donc une diminution de pression, ce qui en retour ralentit la rotation de la roue. Inversement si la vitesse de rotation est anormalement basse, les bras se rapprochent ce qui entraîne une plus grande ouverture des valves se traduisant par une augmentation de pression et par la suite une augmentation de la vitesse. C'est ainsi que le système régulateur de Watt parvient à maintenir une vitesse de rotation constante malgré des changements importants de la pression de vapeur.

Van Gelder présente une solution computationaliste de ce problème. Elle serait un algorithme dont les étapes sont des représentations symboliques depuis l'input (mesure symbolique de la vitesse de la roue) jusqu'à l'output (spécification symbolique de l'action physique sur la valve). Les calculs se succèderaient dans une séquence qui se répète: input / manipulation interne / output. Représentation,

séquentialité et cyclicité sont donc les aspects essentiels de ce modèle computationnel. Or, le régulateur de Watt n'a aucune de ces propriétés. L'angle des bras ne représente pas la vitesse de la rotation. Comme l'angle des bras détermine l'ouverture de la valve et que celle-ci détermine la pression de vapeur qui agit sur la vitesse de la roue, il s'ensuit que la vitesse et l'angle s'influencent mutuellement en tout temps. C'est donc par une interaction mutuelle, simultanée et continue en temps réel que la vitesse et l'angle sont reliés et non par représentation. Seul le langage mathématique de la dynamique peut rendre compte de ce couplage. Il s'agit d'une équation différentielle non-linéaire qui nous donne à chaque instant l'accélération de l'angle des bras en fonction de la valeur de ce même angle, du taux de changement de l'angle et de la vitesse de la roue. La solution de l'équation nous donne la trajectoire du système : l'évolution de l'angle en fonction du temps.

Plusieurs auteurs se sont appliqués à réfuter l'argument non représentationnel de van Gelder (Markman 2000, Eliasmith 1996). Leur démarche consiste à montrer que l'auteur conçoit la représentation dans le sens étroit de symbole. Il n'y a pas de représentation symbolique entre la vitesse de la roue et l'angle des bras. Mais si la représentation est comprise dans son sens présymbolique général, il devient possible d'affirmer un lien de représentation. Considérons la définition largement admise que donne Dretske (1981) de la représentation. Quelque chose représente quelque chose d'autre s'il satisfait deux conditions :

1. Il doit porter de l'information au sujet de cette chose par covariance.
2. Il doit avoir la fonction de porter de l'information au sujet de cette chose.

Cette deuxième condition téléologique permet d'exclure les cas de relations causales qui ne sont pas des représentations. Ainsi la fumée est porteuse d'information sur le feu en ce qu'elle est causée par lui, mais elle n'a pas la fonction de porter cette information. Or il se trouve que le gouverneur de Watt satisfait à ces deux conditions. L'angle entre les bras porte l'information sur la vitesse du moteur puisqu'il covarie

avec elle par une loi de corrélation. Et c'est bien la fonction de l'angle que de porter l'information sur la vitesse puisqu'il a été conçu par Watt dans ce but précis.

Il est donc possible d'analyser un modèle dynamique en termes représentationnels. Tout système de traitement de l'information qu'il soit statique ou dynamique, symbolique ou non symbolique, repose sur une architecture conceptuelle représentationnelle. Les théories qui soutiennent que l'utilisation du concept de représentation ne permet pas une modélisation adéquate des processus cognitifs confondent représentation et symbolicité. En effet, dans le milieu des sciences cognitives on définit fréquemment la représentation comme étant le fait d'être à propos de quelque chose d'autre. Mais cette définition en réalité est celle d'une représentation symbolique. Or la définition classique héritée du Moyen âge est différente. La représentation est la présence originale d'un monde externe dans le monde interne du sujet. Cette définition générale est inévitable parce qu'elle renvoie à la relation entre l'esprit et le monde, entre la pensée et la chose. Que l'on défende une théorie computationnelle ou une théorie dynamique, on parlera toujours de système et d'environnement, ce qui implicitement évoque un lien de représentation.

Le concept de représentation tel que défini tant par les tenants de la computation forte que de la computation molle, ne correspond que peu au concept classique de la représentation. En effet dans presque toute la tradition philosophique, le concept de représentation ne signifie pas "symbole". Dans son sens le plus classique en effet, il signifie simplement *re-présence* d'un monde extérieur dans un système c'est-à-dire présence originale du monde externe dans le monde interne du sujet ou en terme moderne trace ou inscription de l'environnement dans le système. Certes, cette re-présence ou cette inscription peut prendre la forme de symbole. Mais plusieurs autres formes sont aussi possibles. (Meunier, 2002, p.217)

Le concept de représentation est alors une façon commode et même inévitable de parler de certains systèmes en tant qu'ils possèdent en eux une certaine présence de leur environnement et qu'ils en gardent la trace ou l'inscription. Ce que traite le système lui vient de l'extérieur, peu importe la nature de la relation sémiotique avec l'environnement (iconique, causale ou symbolique). Il s'agit de la reconnaissance d'une distinction entre un extérieur et un intérieur. La représentation naît donc à l'intérieur du système, elle est acquise par lui à l'intérieur de lui. Des modèles différents se prononcent sur la nature de cette acquisition. Dans les systèmes artificiels, le mode d'acquisition est de nature conventionnelle alors que dans les systèmes naturels il est de nature causale. Dans les deux cas, le sens fondamental de la représentation comme substitut dans le système de ce qui lui est extérieur est valide. Si ce substitut est un signe créé conventionnellement par un agent interprète, la représentation est dite symbolique, et la relation sémiotique est dite de type sémantique. Si le substitut est né d'un rapport causal entre l'agent et son environnement, elle est dite causale ou indexicale. Mais tous ces modèles soutiennent que l'information résulte d'un rapport quelconque entre le système et l'environnement. Ce que le concept de représentation permet, c'est d'identifier le système par rapport à son environnement et de caractériser de manière générale le processus de traitement d'information interne sans spécifier le mode d'acquisition des données externes. Comme corollaire à cette thèse, toute représentation est idiosyncrasique au système lui-même dans le sens où elle lui est propre. Ainsi face au même environnement la représentation des oiseaux n'est pas la même que celle des chauves souris.

Markman et Dietrich (2000) insistent sur le fait que les sciences cognitives souffrent du fait qu'il n'y a pas une conception générale et acceptée par tous de la représentation.

A significant reason for the debate over representation is that there is no generally accepted definition for the concept.[...] By providing a common framework for talking about representation, we aim to avoid this problem. (Markman & Dietrich, 2000, p.140)

Ces auteurs proposent pour pallier au problème de concevoir la représentation comme un état médiateur (*mediating state*) interne au système, porteur de l'information et qui est utilisé par le système pour atteindre ses objectifs. Selon le phénomène cognitif étudié et selon le modèle utilisé, cet état médiateur se spécifie et se particularise. Il est alors une représentation symbolique, distribué, transitoire, etc. Dans ce cas un état médiateur est une catégorie supérieure plus générale.

4.3. Propriétés de la représentation dans les modèles dynamiques

- *Temporalité*

Ce qui est le plus spécifique au processus cognitif, c'est le temps. Comprendre la dynamique de la cognition, c'est comprendre comment le comportement cognitif change dans le temps. Celui-ci n'est pas comme soutient le modèle computationnel une séquence discrète de représentations à un intervalle arbitraire (t_1, t_2, \dots), mais un système qui évolue de façon continue et simultanée en temps réel. De même, ce ne sont pas l'input et l'output qui ont la primauté dans le processus, mais le changement entre les deux. Le timing est un facteur essentiel de la cognition. Des questions comme le taux, la durée, la synchronie, la période, ... sont importantes et déterminantes. Le représentationnalisme computationnel spécifie la séquence des manipulations symboliques passant d'un état discret à un autre état discret, mais ne dit rien sur le timing de ces états et les transitions entre eux : combien de temps le système va-t-il rester dans le premier état? Quelle est la vitesse de la transition au

deuxième état? Ce sont là des questions importantes ignorées par le paradigme symbolique.

La temporalité inhérente fait de la représentation dynamique un état essentiellement transitoire même si le modèle peut justifier une énergie durable comme c'est le cas de l'attracteur. Cela pose au modèle de réelles difficultés quand il s'agit de produire des représentations suffisamment stables pour rendre compte de certains processus cognitifs de niveau supérieur comme le raisonnement, la généralisation relationnelle, l'utilisation du langage, la reconnaissance d'objets complexes, etc. Non seulement ces processus nécessitent des représentations stables et durables, mais la compositionnalité systématique des représentations s'avère être indispensable dans leur explication.

- *Corporéité*

Une grande innovation de la théorie dynamique est qu'elle exploite les structures de contact avec le monde réel ou les points d'ancrage avec l'environnement qui sont la sensation et l'action. Cela donne aux représentations du modèle un caractère concret qui s'oppose à l'abstraction des symboles dans le modèle classique. Nos représentations internes sont conditionnées par notre incorporation physique et par la nécessité d'une action rapide et solide en réponse à des stimulations sensorielles directes en temps réel. Elles contiennent des informations sur le contexte dans lequel elles étaient formées.

Mais le fait qu'un agent cognitif soit incorporé et enchâssé dans son environnement n'empêche pas qu'il puisse avoir des états représentationnels abstraits. C'est ainsi que la règle d'inférence Modus Ponens est stockée dans l'esprit comme une représentation abstraite et s'applique indépendamment de tout contexte ou de situation extérieure. Nous verrons dans le chapitre suivant comment cette contradiction abstrait /concret disparaît quand on considère la représentation comme

un continuum. La question n'est plus alors de savoir si la représentation est abstraite ou concrète mais celle de mesurer son degré d'abstraction dans un processus cognitif donné.

- *L'obstacle sémantique*

Dire que le modèle dynamique est représentationnel ne règle pas toute la question. Il s'agit de déterminer la fonction de la représentation dans l'économie du modèle et d'identifier ses propriétés. Si on peut dire que le système traite des représentations, ces représentations sont difficilement interprétables. En effet, une difficulté majeure de la théorie des systèmes dynamiques est l'identification des paramètres du système qui compteront dans le modèle.

The central dilemma faced by any experimentalist hoping to apply dynamic systems theory is ignorance, in particular, ignorance of the state variables. (Robertson et al in Eliasmith 1996, p.314)

Comme ce sont des systèmes non linéaires complexes, c'est un défi que d'isoler les paramètres précis et individuels qui les décrivent. Pour contourner cette difficulté, les dynamicistes sont contraints de choisir des paramètres collectifs qui rendent compte du comportement macroscopique global. Le problème, c'est que souvent ces paramètres collectifs ne correspondent à aucune signification précise dans le comportement du système.

Having identified a parameter which controls a macroscopic behavior of an equation describing a system, does not mean that the parameter can be interpreted in the context of the system being described. (Eliasmith 1996, p.315)

C'est là clairement un handicap de la théorie qui la rend moins efficace que les autres. En effet, dans le réseau connexionniste les activations correspondent aux représentations du processus cognitif en question. De même, dans le modèle fonctionnaliste les symboles renvoient clairement à des entités externes. Ce lien sémantique d'interprétation se perd dans la théorie des systèmes dynamiques qui souffre de la difficulté à produire des représentations significatives pour le système. Ce problème fait planer sur la théorie le spectre de l'échec du behaviorisme.

The dynamist assertion that representation is not necessary to adequately explain cognition is strongly reminiscent of the unsuccessful behaviourist project. (Eliasmith 1996, p.318)

Pour conclure, la théorie des systèmes dynamiques n'élimine pas la représentation qui, comprise dans le sens général d'état médiateur porteur d'information, demeure une bonne traduction du couplage dynamique. Mais cela ne sauve pas pour autant la représentation dans ce modèle. Le caractère essentiellement transitoire des états dynamiques ainsi que la généralité et l'arbitraire des paramètres collectifs affectent sérieusement la capacité explicative du modèle. Assurément, la représentation du modèle dynamique n'est pas encore rendue au niveau tant sophistiqué qu'efficace qu'ont atteint la représentation symbolique et la représentation distribuée.

TROISIÈME PARTIE

APPROCHES INTEGRATIVES DE LA REPRÉSENTATION

CHAPITRE V

VERS UNE REPRÉSENTATION MULTIFORME

5.1. 'Remettre la représentation à sa place'

Trois conclusions découlent de ce qui a été dit précédemment.

a) Tous les modèles qu'ils soient symboliques, connexionnistes ou dynamiques, sont représentationnels. Pour peu qu'on s'accorde sur une définition minimale de la représentation comme état médiateur porteur d'information utilisée par le système pour atteindre un objectif, ce concept est efficace et même inévitable pour décrire ces modèles et aussi pour les comparer.

b) Ni le modèle connexionniste ni le modèle dynamique n'ont réussi à supplanter la représentation symbolique réglée du modèle classique. Celle-ci reste encore le meilleur outil explicatif pour certains processus cognitifs de haut niveau.

c) Cependant, les modèles connexionnistes et dynamiques ont élaboré le caractère dynamique de la représentation, en particulier sa sensibilité au contexte, sa temporalité inhérente et sa corporéité. Ces propriétés dynamiques se sont imposées dans l'explication de certains processus de bas niveau. Donc s'ils ne rendent pas caduc le modèle classique, ils démontrent néanmoins son insuffisance.

De ces trois prémisses découlent une conclusion évidente. Il est peu probable que tous les phénomènes cognitifs puissent être expliqués uniquement par l'un de ces

trois modèles. La représentation symbolique conviendra pour certains processus cognitifs tandis que pour d'autres une représentation distribuée ou un couplage dynamique sera plus efficace.

The explananda in psychology and AI [...] range over vastly dissimilar levels of processing. No one representational format can handle all of these levels. The complexity of the cognitive system demands that we use diverse representations, methodologies, and styles of explanation. (Markman & Dietrich 2000, p.162)

Il ne s'agit plus de savoir si la représentation est seulement symbolique ou distribuée ou dynamique, ni de savoir lequel des modèles est le seul qui convient à décrire la cognition, la question est plutôt de savoir quelles propriétés de la représentation sont mobilisées pour expliquer un processus cognitif. Quand on met côte à côte toutes les propriétés que confèrent à la représentation les trois modèles, on est frappé par le fait qu'elles se présentent comme une série de couples bipolaires : discret/continu, durable/transitoire, abstrait/concret, etc. Cela laisse entrevoir la représentation comme un champ de potentialités qui se caractérise par des continuums. Une représentation n'est pas que durable ou transitoire, elle peut avoir une infinité d'états entre les deux.

Il convient donc pour paraphraser Andy Clark de 'remettre la représentation à sa place', de rompre avec l'idée de la représentation comme un concept figé défini qu'il faut prendre ou laisser en bloc. Elle est plutôt à considérer comme un champ de potentiel ou une fonction multidimensionnelle qui peut acquérir des valeurs de durabilité d'abstraction, de continuité, etc. Résolument, la représentation est elle-même un système dynamique.

Representation, properly understood, is a dynamic phenomenon itself, not merely something that may undergo, or be subject to, dynamics. (Bickhard, in Dietrich and Markman 2000, p.31)

Le mérite des modèles dynamiques est de nous avoir montré que la représentation n'est pas que cette entité qui subit un processus, elle est elle-même le fruit d'une dynamique qui préside à sa formation et à sa transformation continue.

There are two broad classes of representational change: processes acting on representations and processes intrinsic to the existence or production of the representations themselves. (Dietrich and Markman 2000, p.7)

Il est intéressant de constater avec Barsotti (2001) que cette double dynamique de la représentation se trouve inscrite dans son sens linguistique. Quand on analyse le mot représentation, on voit que ce concept contient en lui une bipolarité sémantique. En effet, le mot s'applique au *représenter* qui est l'acte de représenter comme dans *je me représente quelque chose* (au moyen d'un représentant mentale). Et il s'applique aussi au *représentant* qui est le médium de représentation comme dans *le mot table représente une planche de bois soutenue par quatre pieds*. On voit facilement ce qui sépare fondamentalement les deux sens. Alors que dans le deuxième sens la représentation est une entité qui participe à un processus dynamique, dans le premier sens elle est elle-même un processus dynamique.

A partir de là, il nous semble que la direction que devraient prendre les recherches est la détermination des mécanismes internes qui président à la formation et à la transformation de la représentation. Deux pistes de recherche nous semblent aller dans ce sens. Ce sont les travaux sur l'encodage et ceux qui portent sur l'explicitation.

- *Encodage / action*

Selon Andy Clark (1996), une représentation interne est supposée jouer deux rôles : encoder l'information sur l'environnement et engager un processus d'action située en

temps réel. Le premier rôle (encodage) est passif et réflexif et le deuxième (contrôle) dynamique et actif. La science cognitive a concentré son intérêt sur l'encodage au détriment du contrôle. Cela explique le caractère trop rigide des modèles classiques et on peut véritablement parler d'une pathologie de la fixation par encodage (*code-fixation*). L'insistance sur l'encodage et l'inférence au détriment de l'action est une caractéristique du modèle classique qui a trop utilisé l'analogie avec la structure syntaxique du langage naturel. Nous devrions procéder à une 'reconceptualisation' de la représentation en accentuant la recherche sur l'action de contrôle et sur l'intervention en temps réel. Pour cela, il faudra tenir en compte la réalité biologique : le contact sensorimoteur avec l'environnement qui a été tant négligé par les modèles classiques.

On trouve chez Bickhard (2000) une position semblable. Selon lui on peut parler d'un encodagisme (*encodigism*) pour décrire le dogme du modèle classique et sa propension à fixer la représentation dans un code interne figé. Certes il y a des représentations qui sont encodées de cette façon et c'est ce qui justifie le traitement symbolique. Mais un tel état d'encodage maximal est une forme particulière. En réalité, la représentation est sujette à une dynamique entre encodagisme et interactionnisme. L'encodagisme est orienté vers le passé. C'est le modèle du spectateur. Il est factuel et non modal en ce qu'il est concerné uniquement par la chaîne causale du passé. Il est explicite sur ce qu'il représente. L'interactionnisme est orienté vers le futur. Il se préoccupe de repérer quels types d'interactions seraient possibles dans une situation particulière. Il est modal et implicite.

- *Représentation explicite / représentation implicite*

L'explicitation de la représentation est postulée par le modèle classique pour justifier l'utilisation des attitudes propositionnelles par la psychologie populaire. En

particulier, le langage de la pensée se présente comme le véhicule syntactique de la représentation explicite.

No intentional causation without explicit representation (Fodor 1987, p.25)

Mais cette spécification est une condition trop sévère qui ne tient pas compte de la réalité des processus cognitifs qui ne font pas toujours appel à une explicitation complète de la représentation. Dans les faits, le degré d'explicitation dépendra de la facilité avec laquelle l'information est récupérée et utilisée. Le système tente d'extraire la signification d'un symbole avec le minimum d'effort. Ainsi pour extraire la signification d'un symbole, le système examine le contexte qui va déterminer la quantité d'effort nécessaire à cette extraction.

The extent to which meaning depends on context determines (in part) the amount of effort involved in recovery of meaning. (Clark in Haugland 2000)

De cette façon on n'envisage plus une explicitation totale et en bloc de la représentation. C'est une approche processive et active qui préside. Nous avons alors affaire à un continuum qui va des formes les plus implicites aux formes les plus explicites. L'effort d'extraction dépendra du contexte. C'est ainsi qu'une représentation qui n'est utilisée que dans un contexte spécifique est plus implicite que celle dont le contenu est exigé pour de nombreuses utilisations.

A full blooded process-oriented account of explicit representation will demand both ease of retrieval and flexibility of use. (Clark in Haugland 2000, p.386)

Le caractère implicite de la représentation trouve sa justification dans le fait que le système soit enchâssé dans son environnement. Un système bien adapté à son environnement contient nécessairement des informations implicites au sujet de cet

environnement. Clark donne l'exemple de la forme d'un poisson qui incorpore implicitement l'information sur l'hydrodynamique de l'eau marine.

5.2. Modèles hybrides

Cette approche d'une représentation comme champ de potentiel à plusieurs variables continues est confirmée par la tendance actuelle de concevoir des modèles hybrides qui intègrent au sein de la même architecture et pour le même processus cognitif plusieurs formes de représentation (Figure 2). On se rend compte de plus en plus que renoncer au pouvoir explicatif de la représentation symbolique dans un modèle connexionniste ou dynamique coûte cher. Pourquoi alors ne pas tenter d'intégrer la représentation symbolique dans ces modèles?

The functional compositionality exhibited by the new wave of connectionist models is proof that at least some degree of compositional structure can be preserved in distributed representations and *readily exploited* (Clark in Haugland 2009, p.389)

Nous présentons deux de ces modèles. Le premier, LISA de Holyoak et Hummel, est un modèle pour le raisonnement analogique qui intègre des représentations symboliques dans une architecture connexionniste. Le deuxième est le Système Symbolique Perceptuel de Prinz et Barsalou qui utilise un modèle classique où les représentations symboliques intègrent les propriétés dynamiques de la sensibilité au contexte, la corporéité et la temporalité inhérente.

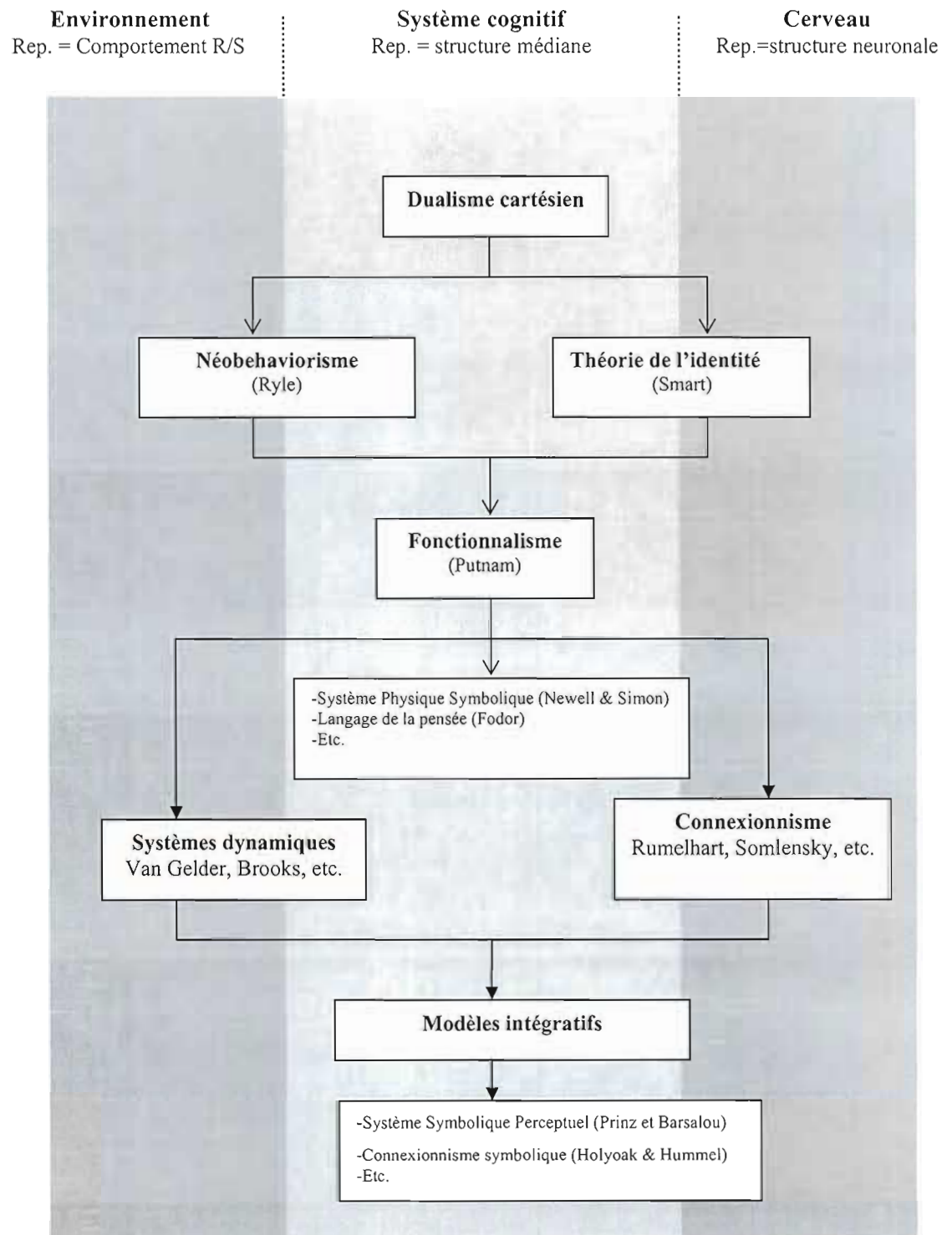


Figure 2. Topographie des modèles cognitifs. (d'après l'auteur)

- *LISA*

Le point de départ de Holyoak et Hummel est le constat l'inévitabilité de l'hypothèse du système physique symbolique.

The human brain support symbol systems ; rather than pretending otherwise, we need to investigate how it does it.[...] given that the PSS hypothesis is accepted and that the brain is apparently a neural network (of some sort), the search for the human cognitive architecture leads in the direction of *symbolic* connectionism. (Holyoak et Hummel in Dietrich et Markman 2000, p.237)

A partir de là, un réseau connexionniste n'a d'autre choix que d'implémenter des représentations symboliques.

What is required, then, is not eliminative connectionism, but rather a proper treatment of symbols in a connectionist architecture, an architecture that simultaneously retains the strengths of distributed representations and instantiates the PSS hypothesis, and hence constitutes symbolic connectionism. (Holyoak et Hummel in Dietrich et Markman 2000, p.231)

C'est une tâche difficile reconnaissent les auteurs car on ne connaît que peu de choses sur la structure neuronale d'une représentation propositionnelle. On connaît cependant les trois conditions nécessaires pour obtenir un réseau symbolico-connexionniste approprié.

a) L'architecture doit être un système symbolique. Ce qui veut dire que le modèle doit fournir des mécanismes pour la composition de symboles et donc pour l'attachement de variable (*variable binding*). Un attachement de variable est un mécanisme qui attribue une valeur à une variable. Il doit être dynamique et

indépendant. Dynamique en ce qu'il peut être créé et détruit tout de suite après. Indépendant en ce que l'attachement ne doit pas dépendre des entités qu'il attache ni changer leur identité. Cette propriété est importante. C'est elle qui explique notre faculté à la généralisation universelle des règles et notre usage de symboles pour intriguer la récupération des structures contenues dans la mémoire.

b) Le réseau doit aussi réaliser des attachements statiques dans la mémoire de long terme. Alors que l'attachement variable dynamique est nécessairement transitoire, une architecture doit aussi établir des attachements statiques pour rendre compte de faits et de règles qui sont durables.

c) Les contenus propositionnels doivent être des représentations distribuées pour rendre compte de la tolérance à l'erreur et de la capacité de récupération basée sur la similitude.

Sur ces bases Holyoak et Hummel ont conçu LISA qui est un modèle connexionniste pour le raisonnement analogique. Le modèle utilise les représentations distribuées mais aussi des représentations symboliques. Le concept est actif quand l'ensemble des nœuds qui le représente a un haut niveau d'activation. La représentation locale est donnée par des nœuds marqués qui correspondent à différents objets et relations. Par exemple le concept John peut être accédé dans sa forme générale en activant son nœud, mais le fait d'activer ce nœud active aussi les nœuds qui lui sont associés. On peut ainsi accéder tant à la représentation locale qu'à la représentation distribuée du concept John dépendamment du contexte. LISA peut à la fois rendre compte d'un contexte particulier et aussi s'abstraire de ce contexte quand cela s'avère nécessaire.

Fait intéressant, la force de ce modèle réside dans ses limites mêmes. En effet, dans les modèles symboliques classiques, l'attachement de variable est libre et peut être fait autant de fois que cela est nécessaire tandis que LISA est limité dans le nombre

d'attachements qu'il peut posséder dans sa mémoire à un temps donné. Mais, soulignent les auteurs, cette limite correspond à la réalité de la cognition humaine qui a une attention et mémoire de travail limitées. Alors que l'approche classique conçoit une intelligence idéale, LISA décrit l'intelligence réelle.

One thing to be gained by asking how the human cognitive architecture implements symbols (rather than simply assuming that it does, as the traditional symbolic approach) is an understanding of some of the limitations of that architecture. (Holyoak et Hummel in Dietrich et Markman 2000, p.268)

- *Le Système Symbolique Perceptuel*

Prinz et Barsalou font le chemin inverse. Ils partent d'une structure symbolique classique et tentent d'y intégrer une composante dynamique.

The central issue is no longer whether internal representations exist, but whether traditional representational approaches can account for the contextually sensitive, embodied, and inherently temporal character of cognition. (Prinz et Barsalou, in Dietrich et Markman 2000, p.51)

Le modèle se base sur la prémisse suivante: les domaines de la connaissance et de la perception ont en commun des mécanismes représentationnels, notamment les zones du sensorimoteur dans le cerveau. Cela est un fait confirmé par la neuroscience qui étudie les pertes de catégories spécifiques. De cela découle une idée intuitive : représenter une catégorie implique l'enclenchement d'un système perceptuel comme si l'agent est en train de percevoir un membre de la classe de cette catégorie. Cela permet donc rompre avec l'approche désincarnée des modèles classiques. Au lieu d'utiliser des symboles amodaux qui sont situés en dehors de la zone perceptuelle, on utilise des représentations dans ces zones mêmes. Le système symbolique perceptuel

peut ainsi créer des représentations qui peuvent à la fois détecter des objets physiques quand ces objets sont utilisés en reconnaissance, et d'avoir un lien symbolique abstrait à ces situations physiques quand elle sont hors contexte.

Perceptual symbol systems attempt to reinvent the strengths of the traditional approaches and dynamic system approach in a single theory. On the one hand, perceptual systems are inherently representational; on the other hand, they are inherently embodied, context sensitive, and temporal. (Prinz et Barsalou, in Dietrich et Markman 2000, p.51)

Le modèle procède par la méthode de la simulation en fournissant un ensemble de mécanismes qui permettent à un individu de simuler les membres d'une catégorie en leur absence. La simulation a deux niveaux de structure. Dans un premier niveau, il y a tout le bagage des connaissances perceptuelle des membres de la catégorie, laquelle expérience est accumulée lors d'une vie comme par exemple la compétence de simulation d'une voiture contient toutes les connaissances sensorimotrices accumulées pendant la vie au sujet des voitures. Au deuxième niveau, la compétence de simulation est capable d'une manière productive de créer des modèles de voitures spécifiques dans différents contextes réels ou virtuels. Donc au lieu de se satisfaire d'une représentation et d'une seule valable pour tous les contextes (ce que fait l'approche classique) la compétence de simulation spécifique crée pour le même objet différentes représentations pour différents contextes. Représenter une voiture en route produit des simulations du point de vue du chauffeur alors que représenter une voiture en réparation produit des simulations de la mécanique des voitures, etc.

CHAPITRE VI

LA DYNAMIQUE REPRÉSENTATIONNELLE CHEZ PIERCE

6.1. L'œuvre de Pierce et la théorie dynamique

Celui qui connaît la théorie des systèmes dynamiques ne manque pas de noter ses multiples résonances chez Peirce dans tous les domaines de sa pensée : catégoriologie, synéchisme, pragmatique, sémiose, fixation des croyances, etc. En effet, pour lui l'explication mécaniciste est surdéterminée et masque la nature dynamique et faillibiliste de la cognition. Peirce conçoit la cognition comme un système dynamique de signes et son œuvre peut être lue comme la recherche d'une conception de la cognition qui serait aussi claire, aussi complète et aussi réussie que la conception de la réalité physique obtenue par la science de la dynamique.

I am not speaking of Soul, the metaphysical substratum of Mind (if it has any), but of Mind phenomenally understood. To get such a conception of Mind, or mental phenomena, as the science of Dynamics affords of Matter, or material events, is a business which can only be accomplished by resolute scientific investigation. (Pierce, CP 7.365)

The law of habit exhibits a striking contrast to all physical laws in the character of its commands. A physical law is absolute. What it requires is an exact relation. [...] On the other hand, no exact conformity is required by the mental law. Nay, exact conformity would instantly crystallize thought and prevent all further formation of habit. The Law of Mind only makes a given feeling more likely to arise. It thus resembles the "non-conservative" forces of physics, such as viscosity and the like,

which are due to statistical uniformities in the chance encounters of trillions of molecules. (CP 6.230)

De plus en plus nombreux sont les études et les recherches qui lisent l'œuvre de Pierce dans une perspective dynamique. Nous en citons ici trois qui nous ont inspiré dans le cadre de cette recherche. Dans son article *Cognition as a Dynamic System*, Richard Tursman soutient que le système de signes se comprend mieux comme un système dynamique. Il analyse comment se combinent les idées en utilisant les concepts de potentiel, de valence et de probabilité.

We may have identified a distinctive pattern [dynamics], perhaps the most distinctive pattern by which the philosophy of Peirce is 'completed' and through which the formulation of a peircean system become explicit (Richard Tursman, 1995)

De son côté, Mark H. Bickhard dans son article *Dynamic Representing and Representational Dynamics* présente son modèle interactionniste comme une alternative à ce qu'il appelle l'encodagisme. Pour lui, le pragmatisme de Pierce est le seul garant historique d'un tel modèle de la cognition.

Cognition is dynamic, and ultimately requires a model in which those dynamics are explained, and not just accommodated. The only historical candidate for accounting for such dynamics is the general pragmatist approach, in which representation and cognition are understood to emerge out of systems of action and interaction. (Bickhard, in Dietrich et Markman 2000, p. 48)

Dans le modèle de Bickhard, la représentation émerge de l'interaction et plus particulièrement de la réponse au problème de sélection de l'interaction appropriée à un environnement particulier.

The pragmatism that Pierce introduced suggests that action and interaction may be essential for understanding consciousness as much as for understanding any other aspect of mind. (p.34)

Enfin citons l'intéressant article de Kelley J. Wells *The Thermodynamic Metaphor. Overdetermination and Peirce's Commitment to Realism*. L'auteur y soutient que la seconde loi de la thermodynamique rend bien compte du système de signes.

We believe that there exists a very powerful interpretive tool by which the subtleties of Peirce's system as he himself practiced it can be rendered more 'clear.' This tool is the thermodynamic metaphor. (Kelley J. Wells, 1997)

6.2. La double fonction du concept de représentation

Dans *Questions Concerning Certain Faculties Claimed for Man* Pierce démontre que toute cognition est nécessairement médiatisée par la représentation. La démarche de Pierce est comme suit. Dans la première question il montre qu'il n'est pas possible de distinguer cognition intuitive et cognition médiatisée. Dans les questions suivantes il soutient qu'il n'est pas possible d'affirmer une auto conscience intuitive, d'établir un pouvoir d'introspection, ou de découvrir une représentation qui n'est pas déterminée par une représentation qui lui précède. Pierce conclut que l'intuition, définie comme 'une cognition non déterminée par une cognition précédente du même objet' (5.213) n'existe pas. La conscience en tant que processus producteur de représentations est la totalité de la relation entre l'esprit et le monde. Toute cognition est nécessairement médiatisée par la représentation mentale. La représentation, dit Peirce, est l'être de la cognition :

Over against any cognition, there is an unknown but knowable reality; but over against all possible cognition, there is only self-contradictory. In short, cognizability

(in its widest sense) and being are not merely metaphysically the same, but are synonymous terms. (CP 5.257)

La sémiotique qui est la «doctrine quasi nécessaire et formelle des signes» se présente comme une science de la représentation au sein de la problématique moderne initiée et ouverte par Descartes. Le signe est universellement déterminé comme une représentation.

Un signe, ou representamen, est quelque chose qui tient lieu pour quelqu'un de quelque chose sous quelque rapport ou à quelque titre. (Écrits sur le signe, Seuil, 1978, p.120, 121)

La représentation a donc une structure triadique : elle est réalisée par un porteur de représentation qui est le signe, elle a un contenu qui est son objet et elle fonctionne comme représentation par un interprète. Si le représentationnalisme classique saisit bien les opérations qui régissent les combinaisons des représentations entre elles, il ne capte pas les changements internes qui président à la formation même de la représentation considérée comme élément statique et inerte. Sensibles au changement et à la variabilité qui sont à l'œuvre dans le phénomène cognitive, les tenants de la théorie dynamique ont en déduit la non pertinence de l'interface représentationnelle et ont opté pour son élimination. Peirce en revanche maintient le processus représentationnel mais élabore sa dynamique.

Dans *On a New List of Categories* (désigné ci-après ONLC), Peirce définit la représentation comme un système dynamique complexe et changeant. Elle est composée de trois catégories premièreté, deuxièmeté et troisièmeité qui renvoient chacune à un élément de la structure triadique. La premièreté se réfère au signe, la deuxièmeté à l'objet et la troisièmeité à l'interprétant. Mais l'originalité de la conception peircienne est que ces catégories ne sont pas des sous-systèmes qui se

partagent de manière complémentaire le système de la représentation. En effet, la totalité de la représentation se confond avec la troisièmeté (*thirdness*) qui intervient dès que le processus cognitif s'enclenche.

We cannot process experience without generalizing. As Peirce says, experience is Thirdness. Consequently the metaphysical (generalization spreading out) is a natural property of experience. (K.J. Wells, P.916)

La représentation est donc d'emblée une troisièmeté qui est la catégorie de la symbolisation. Avec ce biais la conscience procède à l'unification du divers de la substance. La troisièmeté n'adjoint aucun élément nouveau et ne fait que mettre en relation, mettre en ordre le divers de la substance qui est déjà donné.

Thus, the reference to an interpretant arises upon the holding together of diverse impressions, and therefore it does not join a conception to the substance, as the other two references do, but unites directly the manifold of the substance itself. It is, therefore, the last conception in order in passing from being to substance. (CP 1.554)

Or en même temps qu'elle est troisièmeté, la représentation est la totalité des trois catégories. Dans le texte qui suit, Peirce rend bien compte de cette double identité en définissant le signe à la fois comme *whole* et comme *object* :

The very idea of a compound supposes two parts, at least, and a whole, or three objects, at least, in all. (CP 7.537)

C'est la double fonction qui est inhérente au concept de représentation. Elle est d'un côté une symbolisation et se confond avec la troisièmeté, une catégorie-médium. Et de l'autre, elle est une triade d'objets, une totalité où se déploient simultanément les

trois catégories en un processus dynamique. Ailleurs Peirce appelle aussi la troisième catégorie *mise en relation* ou *tiers entre*.

We will call everything marked by being a third or medium of connection, between a first and second anything, *tertian*. (CP 1.297)

Cette propriété complexe de la troisièmeté en tant que tiers est fondamentale et a des conséquences. Elle fonde le statut à double fonction de la représentation et explique à notre avis pourquoi il était si difficile avant Peirce de distinguer représentation et symbolisation et d'abstraire le tiers en tant que concept. Selon ONLC, la conception du tiers est celle d'un objet dont la relation à deux autres est telle que l'un de ces derniers doit se rapporter à l'autre de la même manière que le tiers se rapporte à cet autre. On peut expliciter cette conception comme suit :

- 1- Le tiers est en relation avec le premier et le second. C'est la relation de secondéité du tiers par rapport aux deux autres termes. C'est là la fonction du symbole.
- 2- La relation du tiers au second est la même que la relation du premier au second. Le tiers assure de cette façon une fonction de médiation entre les deux premiers termes. C'est la fonction processus.

En définissant le tiers comme étant à la fois une secondéité et une médiation, Peirce rompt avec le séquentialisme cartésien: le tiers n'est pas un troisième. Même prudence terminologique pour les deux autres catégories : *le soi* est plus approprié qu'*un fondement* car il implique ou appelle l'autre, c'est-à-dire le corrélat. De même, *le second* est plus approprié que *l'autre* car il implique un tiers.

Il me semble donc que voici une doctrine de quelque considérable importance en logique formelle, comme il y a certainement un intérêt pour ceux qui étudient les

trichotomies de montrer dans un cas très simple quelle est l'essence de la trichotomie. Les conceptions de qualité, relation, et mise en relation, commandent ici et ces conceptions apparaissent ici comme identiques à celle de un, autre que, et tiers entre, ou ce qui est presque la même chose aux conceptions de 1, 2, et 3. (in *Écrits sur le signe*, Deladalle, 1978)

6.3. La continuité du champ représentationnel

On voit donc que les catégories sont interdépendantes. En elles-mêmes, elles ne sont pas des entités individuelles qui existent indépendamment les une des autres. Ce sont des variables d'un même système. L'idée aristotélicienne d'une catégorie comme une classe ou un contenant fixe et indépendant est à rejeter. En effet, chez Aristote les catégories sont exclusives et ne s'interfèrent pas. Il n'y a pas par exemple de quantification de la qualité. Au contraire, chez Peirce les catégories ont de sens non seulement par leur signification individuelle mais aussi par leur contribution à la totalité dynamique de la représentation. On peut ici faire appel aux notions de catégories constitutives et catégories agrégatives (E. Mayr in *the growth of biological thought*, 1978). Alors que les catégories aristotéliciennes sont agrégatives, c'est à dire qu'elles sont un outil taxinomique pour classer et découper la réalité, les catégories de Peirce sont constitutives au sens biologique. C'est à dire qu'elles sont des structures qui constituent un organisme ou un système comme par exemple les catégories cellule/tissu/organe ou père/mère/enfant.

Plus importante donc que les catégories en tant qu'individus, est la relation ou le couplage dynamique entre elles. Peirce insiste beaucoup sur cette propriété et il emprunte à la chimie le concept de *valence* comme mesure de dégénérescence des catégories. Dans ONLC l'interdépendance des catégories se manifeste par un concept clé : *la prescission*. Cette méthode d'analyse dynamique consiste à porter l'attention sur un élément et négliger un autre sans toutefois l'ignorer totalement. Elle se situe

entre la discrimination, séparation au sens faible qui consiste à distinguer des termes en fonction de leur sens sémantique, et la dissociation, séparation au sens fort qui consiste dans la conscience d'une chose sans la conscience de l'autre chose. En situant la préscission entre ces deux extrêmes, Peirce évite le piège classique de l'analyse, à savoir la perte d'information. Si nous poussions loin nous dirions que Peirce évite ainsi de tomber dans le piège du dualisme qui dissocie (qui sépare trop) et dans le piège du monisme qui ne fait que discriminer (ne sépare pas assez). Peirce opte pour un principe mitoyen de séparation qui soit souple et dynamique. Le processus consiste à préscinder la conception explicative des conceptions plus immédiates et des impressions. Autrement dit, on peut ignorer ces dernières en gardant la conception qui les a unifiées parce qu'elles sont représentées par l'intermédiaire de la dite conception. Pour cette raison, la préscission est un processus dissymétrique :

Hence, the impressions (or more immediate conceptions) cannot be definitely conceived or attended to, to the neglect of an elementary conception which reduces them to unity. On the other hand, when such a conception has once been obtained, there is, in general, no reason why the premisses which have occasioned it should not be neglected, and therefore the explaining conception may frequently be prescinded from the more immediate ones and from the impressions. (CP 1.549)

La préscission fonde l'articulation dynamique des catégories. Il s'agit de pouvoir les discriminer sans nier leur interdépendance qui est liée à leur degré variable de préscindabilité. Ce processus permet de dégager une identité de la représentation qui prend une forme en fonction de la configuration des catégories. C'est ce qui fait de la représentation un composé hybride, un système dynamique. C'est ainsi que dans le cas du système *signe*, la loi de préscission donne différentes formes de représentation :

- 1- La ressemblance (icône) : c'est quand la premièreté (la qualité) peut être préscindée de la représentation. C'est le cas quand le représentamen et l'objet ont une qualité commune. La représentation est alors iconique et décrit l'expérience phénoménologique et qualitative de la cognition.
- 2- L'indice : c'est quand on peut préscindre la deuxièmeté (la relation) de la représentation. C'est le cas quand le représentamen et l'objet sont liés par un fait. La représentation est alors causale et décrit l'expérience comportementale de la cognition.
- 3- Le symbole : c'est quand rien ne peut être préscindé de la représentation qui trouve sa raison d'être dans l'interprétant qui est le lien conventionnel entre le représentamen et l'objet. La représentation est alors symbolique, elle est l'expérience conceptuelle et généralisatrice de la cognition.

Pour résumer, la représentation est d'emblée une symbolisation qui peut stationner dans le symbole comme elle peut dégénérer en un indice ou un icône. Une fin iconique peut alors donner l'illusion d'un rapport immédiat avec le monde. En réalité elle est le résultat d'un processus intelligible, celui de la prescission à partir de la catégorie de la symbolisation qui est le point de départ. Ainsi le poète ou le mystique ne peuvent prétendre être dans un rapport immédiat avec le monde que par la suite d'un travail considérable sur le symbole (linguistique pour le premier, religieux pour le second) et sa déconstruction.

L'icône, l'indice et le symbole sont en réalité trois familles ou trois classes de formes représentationnelles. En effet, il y a une infinité de représentations. Selon que l'on peut préscindre une qualité ou un certain degré de qualité on obtient une variété d'icônes allant de l'icône pure à l'indice pur, en passant par une infinité d'icônes de plus en plus indiciels. Il est clair par exemple que l'iconicité de la musique est plus

grande que celle du langage, que celle d'une peinture abstraite est plus grande que celle d'une peinture figurative, etc. Assurément on peut attribuer à chaque catégorie une intensité. C'est cette propriété qui fonde la continuité du champ représentationnel. ONLC précise que les conceptions universelles sont les étapes d'un passage graduel de l'indifférencié pur (la substance) à l'unité différenciée. Concept fondamental dans le système peircien, la continuité (ou synéchisme) est une condition de sa dynamique. Déjà la substance est définie par ONLC comme étant un continuum. La conscience n'est jamais dans le vide. Quand elle ne connote pas (c'est-à-dire quand elle ne se représente rien), elle dénote toujours. Et ce qu'elle dénote, c'est le présent indifférencié, le *il*, le donné inanalysé et sans unité propre car préalable à toute synthèse. Comme exemple de substance Peirce cite le x mathématique, la variable continue qui potentiellement admet une infinité de valeurs. Ailleurs, Peirce utilise la métaphore du lac pour illustrer le continuum des représentations (les idées) :

Consciousness is like a bottomless lake in which ideas are suspended at different depths. Indeed, these ideas themselves constitute the very medium of consciousness itself. (CP 7.553)

L'idée est un élément infinitésimal dans un espace continu. Elle a un *degré de liberté* et un *potentiel*, savoir la capacité d'appeler un nombre indéfini d'autres idées.

A continuity obtains between any idea and any other idea by virtue of the circumstance that every idea is a differential element of the same potential field. (Tursman, 1995)

Le champ représentationnel est continu parce que les catégories peuvent se combiner pour donner lieu à une infinité de configurations possibles. La représentation est un système dynamique qui selon l'intensité et la combinaison de ses paramètres (les

catégories), peut prendre toutes les formes possibles dans un espace continu. Chaque point de cet espace est une représentation particulière. Nous dirions que la fonction *représentation* acquiert en ce point une valeur.

6.4. L'évolution du processus représentationnel dans le temps

Comme système dynamique, la représentation a une histoire et évolue dans le temps. Nous croyons avec K. J. Wells que la métaphore thermodynamique est une approche très efficace pour décrire cette évolution. En effet, le système catégoriel (la fonction *représentation*) produit un travail thermodynamique dans la forme de la troisièmeté, lequel travail consiste concrètement dans le phénomène de la généralisation. Parallèlement à cette production positive et comme condition nécessaire de celle-ci, la représentation génère aussi l'entropie sous forme de deuxièmeté. Autrement dit, pour produire la troisièmeté, il faut payer le prix de dépenser de la deuxièmeté. Ceci se comprend bien par exemple dans le cas de la représentation scientifique. On sait que pour établir une loi (troisièmeté), il faut effectuer de nombreuses expériences empiriques (deuxièmeté) dont seulement une partie sera réellement utile. Ainsi la représentation tend naturellement à la généralisation. C'est le phénomène de la propagation des idées qui se fait selon la loi de l'esprit (Law of Mind) qui est une loi incontestablement dynamique:

Logical analysis applied to mental phenomena shows that there is but one law of mind, namely, that ideas tend to spread continuously and to affect certain others which stand to them in a peculiar relation of affectability. In this spreading they lose intensity, and especially the power of affecting others, but gain generality and become welded with other ideas. (CP 6.104)

Le système sémiotique des idées tend vers l'équilibre par le biais du processus de la généralisation. Plus une idée gagne en généralité plus elle perd en intensité, la

tendance étant d'atteindre une généralisation maximale. Au début de ce processus la représentation s'appuie sur la deuxième. Mais plus la troisième gagne du terrain, plus elle veut s'affranchir de la dépense coûteuse de la deuxième, le travail thermodynamique décline en faveur d'une inertie croissante jusqu'au moment où la généralisation atteint son maximum : la troisième règne alors sans partage et la deuxième n'intervient plus. C'est alors la mort de la représentation qui ne se nourrit plus de l'expérience et de la lutte. Seule alors une intervention de la première, source de créativité et de chance, peut relancer de nouveau la dynamique de la représentation.

En tant que système thermodynamique, la représentation s'affaiblit (*run down*) en cherchant l'équilibre. En se propageant elle gagne en généralité et perd en intensité pour enfin se stabiliser dans une forme symbolique de généralisation. L'analogie est à faire avec la chaleur qui se dissipe et cherche une uniformisation de température (équilibre) à l'intérieur d'un système thermodynamique clos. C'est la seconde loi de la thermodynamique qui est à l'œuvre. Les idées sont comparées aux non-uniformités de la température.

Quand la représentation atteint l'équilibre du symbole il n'y plus de travail thermodynamique à faire. Il y a quelque chose de mort dans la vie de la représentation qui s'éteint ainsi. Autrement dit, la représentation ne peut plus faire un travail philosophique.

Thus an idea fully spread out, a generalization covering everything has no intensity. Consequently, it has no utility or ability to do philosophical work. (K.J.Wells, p.902)

En effet, quand la représentation se propage totalement, la signification ne signifie plus rien, la généralité ne généralise plus, il y a quelque chose qui meurt dans le processus, la loi est arbitraire. La vie meurt, l'esprit est inintelligible. Mais par chance

(techysm) une nouvelle loi ou une nouvelle généralisation peut émerger et relancer le processus. Cette chance, c'est la catégorie de la premièreté qui en est le lieu.

L'énergie qui contribue au travail mécanique d'un moteur a comme équivalent dans la loi de l'esprit l'énergie médiatisée ou généralisée. Elle est reliée à un objectif supérieur qui la transforme en travail. Cette énergie prend la forme de la troisième. Donc qualitativement l'énergie en forme de troisième est supérieure à l'énergie gaspillée ou dépensée qui prend la forme de deuxième. Dit autrement, la deuxième est l'énergie que le système dépense pour élaborer des généralisations sous forme de troisième.

Ainsi dans le domaine de la croyance et du doute, c'est le même processus thermodynamique qui est à l'oeuvre. Toute croyance particulière est le produit d'un spectre plus large de croyances. Aucune croyance singulière ne peut se soutenir elle-même indépendamment d'autres croyances non soutenues. Par conséquence, toutes les croyances consomment plus d'énergie en forme de croyances non soutenues qu'elles ne sont de libres et autonomes croyances.

CONCLUSION

Dès leur apparition dans le champ du savoir, les sciences cognitives ont postulé la représentation mentale pour rendre compte de plusieurs aspects de la cognition humaine et animale. Pendant longtemps a dominé la thèse du fonctionnalisme computationnel qui conçoit la représentation comme symbole. Dès les années 80 des théories rivales ont fortement critiqué ce modèle et ont radicalement opté pour le rejet du symbole, voire même de la représentation, au profit d'une approche dynamique de la cognition.

Les dernières dix années voient les choses se replacer. L'obsession paradigmatique qu'ont eu successivement les modèles symboliques, connexionnistes et dynamiques s'estompe. Et c'est le concept de représentation qui sort gagnant de ces débats. Son rôle s'avère encore plus complexe que ses premiers défenseurs ne l'auraient imaginé. Non seulement la représentation participe à un processus qui lui est extérieure, elle est elle-même le fruit d'une dynamique qui lui est interne et qui préside à sa formation et sa transformation. Les premiers modèles cognitifs qui considéraient la représentation comme une entité solide et indivise, ont investi tous les efforts pour analyser le processus externe. Nombreux sont les modèles aujourd'hui qui s'aventurent 'à l'intérieur' de la représentation. Des concepts naissent qui tentent de caractériser cette dynamique indépendamment du modèle postulé. L'encodage et l'explicitation sont deux parmi ces concepts.

L'œuvre de Peirce apporte un éclairage puissant dans l'étude de la représentation. Elle initie à notre avis un mouvement d'intégration et de synthèse entre symbolisme et dynamisme dont la portée pourrait être comparable à celle qu'a opérée la critique kantienne entre rationalisme et empirisme. Selon Peirce, il n'est pas possible pour l'esprit humain d'être en rapport avec le monde par d'autre biais que la

représentation. N'en déplaise aux nostalgiques d'une certaine phénoménologie mystique qui prône l'immédiateté de l'expérience, le rapport de l'homme au monde est toujours médiatisée par la représentation. Dès que la conscience veut concevoir le divers de la substance, la catégorie de la symbolisation, la troisième, intervient comme système de configuration de trois catégories. Par un effort intelligible l'esprit peut déconstruire la représentation triadique et la faire dégénérer à ses composantes indexicales ou iconiques.

Dès lors, le conflit qui se pose n'est pas celui entre une conception qui serait représentationnelle et une autre qui serait anti-représentationnelle mais, au sein même de la représentation, entre dynamisme et symbolisme. Déclenchée dès le départ comme un processus de symbolisation, le système-représentation tend à stationner dans le symbole mais la dynamique de la deuxième qui est l'objet de la représentation et surtout de la première qui est le facteur d'intuition et de chance tentent de résister à la fixation symbolique et de relancer le processus. Il y a à l'intérieur de la représentation une tension entre troisième et première qui est la tension entre phénoménologie et symbolisation, entre conception mystique et conception technique du monde. Ce conflit est dans la nature même de la cognition. La troisième est le pôle de la technique, elle nous donne des instruments pour lutter contre la menace extérieure et assurer notre survie, mais elle nous fait payer le prix cher : en agissant ainsi nous courons le risque de nous couper du socle naturel. Et en se coupant de la nature nous nous coupons de notre propre créativité. Une troisième poussée à l'extrême conduit à la sécheresse de l'être, à la mort. D'où l'association entre représentation et mort dans certains courants de la philosophie continentale (Heidegger, Foucault, Artaud). L'homme a besoin de la fraîcheur de la première qui l'inscrit dans son environnement naturel et dans les origines, mais la première seule ne peut pas assurer la survie de l'homme qui est en perpétuelle lutte pour s'adapter à un environnement hostile.

Prenons l'exemple du langage. Ce dernier est l'outil représentationnel le plus complexe de l'esprit humain. Contrairement à l'animal qui est contraint d'inscrire sa cognition dans sa biologie, l'homme dépose ses moyens de survie dans le support symbolique du langage. Cela lui permet d'enrichir et de complexifier ces moyens et aussi d'assurer leur transmission. Mais ce faisant, l'homme paie le prix de s'aliéner dans sa culture au point de s'arracher au socle naturel. C'est la fonction de l'art que d'aller dans le sens contraire de la technique. Le technicien doit déjouer l'animalité et renforcer la représentation du langage, de l'instrument. L'artiste doit au contraire nous ramener constamment à notre nature première en déconstruisant nos représentations, en les remettant toujours en question, en renouvelant notre alliance mystique à l'univers. Ainsi si le mot est une représentation symbolique qui trouve sa raison d'être dans la convention langagière, la poésie va s'efforcer de le rendre un icône qui existe par lui-même. Au sein du poème, le mot échappe à sa définition linguistique conventionnelle et existe comme une qualité sonore, voire même, dans certains langues, comme une qualité picturale (calligraphie). Mais la poésie seule ne peut pas assurer la survie de l'homme qui a besoin de maîtriser son environnement. Ainsi le débat philosophique au sujet de la représentation se trouve ramené grâce à cette approche à sa vraie nature. La représentation est une dynamique complexe. La cognition humaine se nourrit de cette dynamique. Elle a autant besoin de la force symbolique de la troisième que de l'énergie phénoménologique de la première. La première lui confère la maîtrise de la technique, la seconde lui donne la créativité de l'art.

Nous le savons, la représentation est une question philosophique qui dépasse les limites des sciences cognitives. Elle traverse tous les champs du savoir et de l'action : sciences sociales, arts, théologie, etc. C'est une question profonde qui renvoie aux antagonismes qui trouvent leur racine dans un vieux conflit de la philosophie occidentale. Il s'agit de ce que Charles Taylor a désigné comme «la guerre globale entre les héritiers des Lumières et les Romantiques» dont l'enjeu est la nature de

l'homme, ou encore la question de savoir ce que c'est que être humain, ce que c'est que penser et créer. Ainsi selon Bernard Barsotti :

La philosophie a toujours été, et la philosophie contemporaine l'est plus que jamais, le lieu d'une guerre de la représentation, dont l'enjeu final, parfois oublié derrière des débats excessivement techniques, est la valeur même de l'humanité, à travers l'évaluation de son pouvoir de vérité ou d'accès à la réalité vraie, ainsi que le statut métaphysique de l'homme, à travers l'évaluation de la possibilité de se réaliser (ou non) qui est réservée à son désir de surmonter sa finitude. (Barsotti, p.6)

Il n'est pas inutile de dire à la fin de ce mémoire un mot sur ce qui m'a amené à cette recherche. Je suis venu à la réflexion philosophique depuis l'écriture théâtrale. Le théâtre comme chacun sait, ce n'est rien d'autre que l'art de la représentation. Venant d'une culture arabo-musulmane, je suis confronté à une question qui a fait coulé beaucoup d'encre et a tordu le cou à beaucoup de chercheurs. Les arabes qui, au moyen âge ont traduit et commenté de fond en comble tout le corpus de la philosophie, des mathématiques et de la médecine de la Grèce antique, n'ont pas eu le moindre intérêt pour son théâtre. Pas une seule tragédie n'a été traduite ni même signalée dans la littérature².

L'islam adopte à l'égard de la représentation une attitude de méfiance et d'hostilité. Il n'existe pas de clergé supposé représenter et médiatiser dieu. La représentation de dieu et du vivant par le langage a été strictement prohibée. C'est ce qui explique que l'art musulman n'a connu ni peinture ni sculpture figuratives. La littérature arabe qui a atteint des sommets inégalables dans la poésie et la prose mystique s'est interdite

² L'écrivain argentin Borges a exploité admirablement cela dans sa nouvelle *La quête d'Averroès*. On y voit le philosophe andalou en train de traduire *La poétique* d'Aristote; il achoppe sur les mots

toutes les formes de la fable et de la fiction (*Les milles et une nuits* étant une exception, d'origine indo-persane du reste). Aujourd'hui, ils sont nombreux les analystes qui expliquent la difficulté voire même l'incapacité des sociétés musulmanes à installer une véritable démocratie parlementaire par le refus de la représentation politique.

Il est indéniable que l'islam a une vision dynamique et interactionniste du monde pour laquelle nombre de penseurs de la mouvance dite postmoderne dissimulent mal leur admiration (tout le monde se souvient de l'enthousiasme de Foucault pour Khomeiny et la révolution islamique). Artaud a trouvé dans les confréries soufies marocaines l'inspiration pour son théâtre de la non représentation. Matisse et Picasso se sont beaucoup inspirés des arts musulmans (architecture, tapisserie, céramique, calligraphie...) qui ont démontré que l'art plastique peut se libérer complètement de la figuration pour devenir un déploiement iconique et libre des formes et des couleurs. Tout cela est vrai. Mais il est grand temps que les penseurs musulmans intègrent quelque chose du représentationnel de l'occident moderne, ce qui me renvoie à ma propre responsabilité d'artiste et d'intellectuel.

Quand à la philosophie occidentale, elle fait face aujourd'hui à des grands défis. La globalisation du monde la mettra de plus en plus en choc avec d'autres visions du monde. Cela se présentera à elle tant comme une promesse d'enrichissement qu'une réelle menace qui pèsera sur ses acquis chèrement conquis. Depuis Heidegger, la mode est de flirter avec les philosophies orientales admirées pour leur non représentationnel. La représentation symbolique est un des acquis de la modernité. On peut dire sans exagérer que sa naissance, son développement et ses apports sont à mettre en parallèle avec ceux du rationalisme. Qu'il faille compléter et complexifier

tragédie et comédie qu'il ne sait pas traduire en arabe ne sachant pas ce qu'ils signifient et ne possédant aucune référence de la représentation théâtrale.

la représentation à la lumière des critiques venues d'autres horizons, cela est certain et nécessaire. Mais gare à la tentation de jeter le bébé avec l'eau du bain !

En quoi ces grandes questions concernent ce mémoire aussi petit que maladroit qui s'intitule *Représentation et symbole dans les sciences cognitives* ? Je fais le pari que les sciences cognitives seront le lieu stratégique pour le combat de la représentation parce qu'elles seules sont qualifiées pour être un intermédiaire crédible entre les sciences de la nature et les sciences sociales. Je suis Dan Sperber³ dans son hypothèse que les sciences sociales n'ont d'autre choix que d'adopter le matérialisme minimal en s'appuyant pour cela sur les sciences cognitives. A ce jour, ces dernières sont les seules à pouvoir fournir l'outil scientifique qui permettra à la représentation de continuer à fleurir en intégrant les critiques de ses adversaires sans jamais succomber à leur anathème.

On jugera donc la représentation à l'aune de ses évolutions possibles : loin d'être un principe figée elle s'affine, se complexifie, s'invente, dans tous les domaines, de nouvelles stratégies et de nouvelles règles, cherchant toujours mieux à amener l'idée à la figure, le projet au programme, la force à l'expression. Dans la représentation scientifique, on décèlera tout autant la mise à plat brutale objectivante des qualités sensibles de la chose, que l'effort des mathématiques récentes pour intégrer de mieux en mieux les spécificités qualitatives du phénomène dans la connaissance. Dans la représentation politique, on verra tout autant la mise à plat et la neutralisation des forces vitales en présence dans la forme parlementaire, que la lutte pour l'accès à de nouvelles formes modernes de la représentativité. (Barsotti, 2001, p.15-16)

³ Nous renvoyons à son article *Les sciences cognitives, les sciences sociales et le matérialisme* dans Andler (1993).

BIBLIOGRAPHIE

- Andler, D. 1993. *Introduction aux sciences cognitives*. Paris : Gallimard.
- Barsotti, Bernard. 2001. *La représentation dans la philosophie contemporaine*. Paris : ellipses.
- Brooks, R.A. 1997. «Intelligence without representation». *Artificial Intelligence* 47: 139-159. In J. Haugeland, Ed., *Mind Design II: Philosophy, Psychology, Artificial Intelligence*. 2nd ed., revised and enlarged. Cambridge, MA: MIT Press, pp. 395–420.
- Churchland, P. M. 1989. *The Neurocomputational Perspectives*. Cambridge: MIT Press.
- Clark, A. 1997. *Being There: Putting the Brain Body and World together Again*. Cambridge: Cambridge UP.
- Cummins, R. 1989. *Meaning and Mental Representation*. Cambridge: MIT Press.
- Dennett, D. 1987. *Intentional Stance*. MIT Press
- Dennett D. 1996. *Kinds of Minds*. New York : Brockman
- De Tienne, André. 1996. *L'analytique de la représentation chez Peirce. La genèse de la théorie des catégories*. Bruxelles : Facultés universitaires Saint-Louis.
- Dietrich E. et Markman A. 2000. *Cognitive Dynamics: Conceptual and Representational Change in Humans and Machines*.
- Dreyfus, Hubert. 1998 Intelligence without representation. In <http://www.hfac.uh.edu/cogsci/dreyfus.html>
- Eliasmith, C. (1996). The third contender: A critical examination of the dynamicist theory of cognition. *Philosophical Psychology*, 9, 441-463.
- Fisette, Jean. 1990. *Introduction à la sémiotique de C. S. Peirce*. Montréal : XYZ.
- Flanagan, O. 1992. *Consciousness Reconsidered*. Cambridge: MIT press.
- Fodor, J. A. 1975. *The Language of Thought*. New York: Crowell.
- Fodor, J. A. 1981. *Representations*. Cambridge: MIT Press.
- Fodor, J.A. et Pylyshyn, Z.W. 1988. *Connectionism and Cognitive Architecture: A Critical Analysis*. *Cognition*, 28(1-2), 3-71.
- Freeman ,W. 1991. «The physiology of perception». *Scientific American*. 264-78-85.
- Gibson, J. J. 1950. *The Perception of the Visual World*. Boston: Houghton Mifflin.

- Haugeland, J. 1986. *Artificial Intelligence: The very Idea*. Cambridge: Bradford Book MIT Press.
- Haugeland, J. 2000. *Mind Design II*. Cambridge: Bradford Book MIT Press.
- Jackendoff, R. 1987. *Consciousness and the Computational Mind*. Cambridge: Bradford Book MIT Press.
- Johnson-Laird, P. N. 1988. *The Computer and the Mind*. Harvard UP.
- Markman A. et Dietrich E. 2000. «In Defense of Representation» in *Cognitive Psychology* 40, 138-171.
- Marr, D. 1982. *Vision: A computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman.
- Maturana, H.R et Varela, F. J. 1980. *Autopoeisis and Cognition: The Realization of the Living*. Dordrecht, Boston, London, D Reidel.
- Meunier, J.G. 2000. «Niveau de représentation et information» in F" Braga " La representation " U.de Trento: sous presses.
- Meunier, J. G. 2002. «Représentation et non computationnalité» *Representazione e teorie della mente*. P. Univ. Parma. p. 211-229.
- Newell, A. et Simon, H. 1976. «Symbol Manipulation». In *Encyclopedia of Computer Science*. New York: Petrocelli/Charter.
- Peirce, Charles. S. 1931-1958. *Collected Papers*. Cambridge: Harvard University Press.
- Peirce, Charles S. 1978. *Écrits sur le signe*. Rassemblés, traduits et commentés par Gérard Deladalle. Paris : Editions du Seuil.
- Penrose, R 1994. *Shadows of the mind*. Oxford: Oxford U. P.
- Petitot, J. et Varela, F. 1999. *Naturalizing Phenomenology*. California : Stanford University Press.
- Port, R. et Van Gelder, T. 1995. *Mind as Motion*. Cambridge: MIT Press.
- Putnam, H. 1988. *Representation and Reality*. Cambridge: MIT Press.
- Reynolds, Andrew. 1996. «Peirce's Cosmology and the Laws of thermodynamics». *Transaction of the Charles S. Peirce Society*, vol. 32, no 3 (été).
- Rosensohn, William L. 1974. *The phenomenology of Charles S. Peirce*. Amsterdam: B.R. Grüner B.V.
- Rumelhart, D. et Mc Lelland, J. L. 1987. *Parrallel Distributed Processing*. Cambridge:

MIT Press.

Smolensky, P. 1988. «On the Proper Treatment of Connectionism». *The Behavioral and Brain Sciences*, II 1-74

Sperber, D. 1974. *Le symbolisme en général*. Paris: Herman.

Thelen, E. et Smith, L.B. 1994. *A dynamic systems approach to the development of cognition and action*. Cambridge, MIT Press.

Tursman, Richard. 1995. «Cognition as a dynamic system». *Transaction of the Charles S. Peirce Society*, vol XXXI, no 2 (automne)

Varela, F. J. 1988. *Invitation aux sciences cognitives*. Paris: Seuil.

Van Gelder, T. 1997. «Dynamics of Cognition» In Haugeland, (ed). J(1997) *Mind design II*, MIT Press.

Wells Kelley, J. 1997. «The Thermodynamic Metaphor. Overdetermination and Peirce's Commitment to Realism» *Transaction of the Charles S. Peirce Society*, vol XXIII, no 4 (automne)